

09/171,432

DIALOG

Set	Items	Description
S1	328	AU="FIELDS H" OR AU="FIELDS H A"
S2	298	AU="FIELDS H.A." OR AU="FIELDS HA"
S3	75	AU="FIELDS HOWARD" OR AU="FIELDS HOWARD A"
S4	175	AU="KHUDYAKOV Y" OR AU="KHUDYAKOV Y E" OR AU="KHUDYAKOV Y.- E." OR AU="KHUDYAKOV YE"
S5	49	AU="KHUDYAKOV YURY" OR AU="KHUDYAKOV YURY E"
S6	733	S1 OR S2 OR S3 OR S4 OR S5
S7	19056	HEPATITIS (W) A (W) VIRUS
S8	14247	HAV
S9	26065	S7 OR S8
S10	543159	VACCINE? ? OR IMMUNOGENIC
S11	1376	P2A
S12	106	S6 AND S7
S13	88	S12 NOT PY>1997
S14	42	RD (unique items)
S15	1	S11 AND S14
S16	10	S14 AND S10
S17	18	S9 AND S10 AND S11
S18	18	S17 NOT PY>2996
S19	6	S18 NOT PY>1997
S20	6	RD (unique items)
?		

THIS PAGE BLANK (USPTO)

16/3,AB/1 (Item 1 from file: 155)
 DIALOG(R) File 155:MEDLINE(R)
 (c) format only 2000 Dialog Corporation. All rts. reserv.

06034651 86200371

Structure of the hepatitis A virion: peptide mapping of the capsid region.

Wheeler CM; Robertson BH; Van Nest G; Dina D; Bradley DW; Fields HA
 Journal of virology (UNITED STATES) May 1986, 58 (2) p307-13, ISSN
 0022-538X Journal Code: KCV
 Languages: ENGLISH
 Document type: JOURNAL ARTICLE

Milligram amounts of highly purified hepatitis A virus (HAV) were obtained from persistently infected cell cultures. The HAV polypeptides were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis and transferred to nitrocellulose for detection by an enzyme-linked immunotransfer blot procedure. The HAV nucleotide-derived amino acid sequence was subjected to computer analysis to identify potential immunogenic regions within the HAV capsid polypeptides. Synthetic peptides corresponding to selected regions of each of the larger putative capsid polypeptides were coupled to keyhole limpet hemocyanin and used to immunize rabbits. Four of six anti-HAV peptide sera were strongly reactive. Antipeptide serum generated against amino acids (a.a.) 75 through 82 reacted with the 27,000-molecular-weight (MW) polypeptide; serum against a.a. 279 through 285 reacted with the 29,000-MW HAV polypeptide; and sera against a.a. 591 through 602 and 606 through 618 reacted with the 33,000-MW HAV polypeptide. These reactions enabled the identification of the gene order of the larger HAV P1 region gene products. Our data indicate the following molecular weights: HAV VP2 or 1B, 27,000; HAV VP3 or 1C, 29,000; and HAV VP1 or 1D, 33,000.

16/3,AB/3 (Item 1 from file: 357)
 DIALOG(R) File 357:Derwent Biotechnology Abs
 (c) 2001 Derwent Publ Ltd. All rts. reserv.

0219271 DBA Accession No.: 98-00868 PATENT

Immunogenic hepatitis A virus peptides- recombinant peptide and
 DNA sequence

AUTHOR: Fields H A ; Khudyakov Y E

CORPORATE SOURCE: Atlanta, GA, USA.

PATENT ASSIGNEE: U.S.Govt.; U.S.Dep.Health-Hum.Serv.Atlanta 1997

PATENT NUMBER: WO 9740147 PATENT DATE: 971030 WPI ACCESSION NO.:
 97-535831 (9749)

PRIORITY APPLIC. NO.: US 15644 APPLIC. DATE: 960419

NATIONAL APPLIC. NO.: WO 97US6891 APPLIC. DATE: 970418

LANGUAGE: English

ABSTRACT: New peptides (72 protein sequences specified) from immunogenic hepatitis A virus (HAV) may be used to induce an immune response to HAV in a mammal or to detect the presence of antibodies against HAV in a mammal. DNA sequences encoding the proteins are also claimed, and may be used to produce the recombinant peptides. Some of the claimed peptides can be used to differentiate between vaccine-induced immunity and natural immunity. The peptides can also be used to raise antibodies. The peptides are preferably conjugated to a carrier protein chosen from serum albumin, keyhole limpet hemocyanin, diphtheria toxin, tetanus toxin and synthetic polymers. (140pp)

16/3,AB/4 (Item 1 from file: 342)

THIS PAGE BLANK (USPTO)

DIALOG(R)File 342:Derwent Patents Citation Indx
(c) 2001 Derwent Info Ltd. All rts. reserv.

02989838 WPI Acc No: 97-535831/49

Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune response to HAV in a mammal or to detect the presence of antibodies against HAV in a mammal

Patent Assignee: (USSH) US DEPT HEALTH & HUMAN SERVICES

Author (Inventor): **FIELDS H A ; KHUDYAKOV Y E**

Patent (basic)

Patent No	Kind Date	Examiner Field of Search
WO 9740147	A1 971030 (BASIC)	

Derwent Week (Basic): 9749

Priority Data: US 15644 (960419)

Applications: AU 9728103 (970418); WO 97US6891 (970418)

Designated States

(National): AU; CA; JP; US

(Regional): AT; BE; CH; DE; DK; ES; FI; FR; GB; GR; IE; IT; LU; MC; NL; PT; SE

Derwent Class: B04; D16; S03

Int Pat Class: A61K-039/12; C07K-014/10; C12N-015/11

Number of Patents: 002

Number of Countries: 021

Number of Cited Patents: 003

Number of Cited Literature References: 003

Number of Citing Patents: 000

16/3,AB/5 (Item 1 from file: 349)

DIALOG(R)File 349:PCT Fulltext

(c) 2001 WIPO/MicroPat. All rts. reserv.

00538726

**ANTIGENICALLY REACTIVE REGIONS OF THE HEPATITIS A VIRUS POLYPROTEIN
REGIONS REACTIVES AUX ANTIGENES DE LA POLYPROTEINE DU VIRUS DE L'HEPATITE A**

Patent Applicant/Assignee:

THE GOVERNMENT OF THE UNITED STATES OF AMERICA represented by THE
SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES Centers for
Disease Control and Prevention Technology Transfer Office, THE
GOVERNMENT OF THE UNITED STATES OF AMERICA, represented by THE
SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES, Centers for
Disease Control and Prevention, Technology Transfer Office , Executive
Park, Building 4, Suite 1103, MS E­67, Atlanta, GA 30329 , US

Inventor(s):

FIELDS Howard A , FIELDS, Howard, A. , 1823 Jackson Circle Drive,
Marietta, GA 30333 , US

KHUDYAKOV Yury E , KHUDYAKOV, Yury E. , 4354 Tremont Court, Duluth, GA
30136 , US

Patent and Priority Information (Country, Number, Date):

Patent: WO 9740147 A1 19971030

Application: WO 97US6891 19970418 (PCT/WO US9706891)

Priority Application: US 9615644 19960419

Designated States: AU CA JP US AT BE CH DE DK ES FI FR GB GR IE IT LU MC NL
PT SE

Publication Language: English

Filing Language: English

Fulltext Word Count: 33399

English Abstract

The present provides **immunogenic** HAV peptides, antibodies and assays
for detecting HAV. **Vaccines** against HAV are also provided.

THIS PAGE BLANK (USPTO)

French Abstract

Cette invention concerne des peptides du virus de l'hepatite A (VHA)
immunogenes, des anticorps et des dosages permettant de detecter le VHA.
Des vaccins diriges contre le VHA sont egalement presentes.

?

THIS PAGE BLANK (USPTO)

20/3,AB/3 (Item 1 from file: 399)
 DIALOG(R) File 399:CA SEARCH(R)
 (c) 2001 AMERICAN CHEMICAL SOCIETY. All rts. reserv.

128012554 CA: 128(2)12554h PATENT
Immunogenic hepatitis A virus peptides, vaccines and antibodies and assays for detection of virus
 INVENTOR(AUTHOR): Fields, Howard A.; Khudyakov, Yury E.
 LOCATION: USA
 ASSIGNEE: United States Dept. of Health and Human Services; Fields, Howard A.; Khudyakov, Yury E.
 PATENT: PCT International ; WO 9740147 A1 DATE: 19971030
 APPLICATION: WO 97US6891 (19970418) *US 15644 (19960419)
 PAGES: 139 pp. CODEN: PIXXD2 LANGUAGE: English CLASS: C12N-015/11A; C07K-014/10B; A61K-039/12B; G01N-033/576B DESIGNATED COUNTRIES: AU; CA; JP ; US DESIGNATED REGIONAL: AT; BE; CH; DE; DK; ES; FI; FR; GB; GR; IE; IT; LU; MC; NL; PT; SE

20/3,AB/4 (Item 1 from file: 349)
 DIALOG(R) File 349:PCT Fulltext
 (c) 2001 WIPO/MicroPat. All rts. reserv.

00538726

**ANTIGENICALLY REACTIVE REGIONS OF THE HEPATITIS A VIRUS POLYPROTEIN
 REGIONS REACTIVES AUX ANTIGENES DE LA POLYPROTEINE DU VIRUS DE L'HEPATITE A**
 Patent Applicant/Assignee:

THE GOVERNMENT OF THE UNITED STATES OF AMERICA represented by THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES Centers for Disease Control and Prevention Technology Transfer Office, THE GOVERNMENT OF THE UNITED STATES OF AMERICA, represented by THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES, Centers for Disease Control and Prevention, Technology Transfer Office , Executive Park, Building 4, Suite 1103, MS E­67, Atlanta, GA 30329 , US

Inventor(s):

FIELDS Howard A, FIELDS, Howard, A. , 1823 Jackson Circle Drive, Marietta, GA 30333 , US
 KHUDYAKOV Yury E, KHUDYAKOV, Yury E. , 4354 Tremont Court, Duluth, GA 30136 , US

Patent and Priority Information (Country, Number, Date):

Patent: WO 9740147 A1 19971030
 Application: WO 97US6891 19970418 (PCT/WO US9706891)
 Priority Application: US 9615644 19960419

Designated States: AU CA JP US AT BE CH DE DK ES FI FR GB GR IE IT LU MC NL PT SE

Publication Language: English

Filing Language: English

Fulltext Word Count: 33399

English Abstract

The present provides **immunogenic HAV** peptides, antibodies and assays for detecting **HAV** . **Vaccines** against **HAV** are also provided.

French Abstract

Cette invention concerne des peptides du virus de l'hepatite A (VHA) immunogenes, des anticorps et des dosages permettant de detecter le VHA. Des vaccins diriges contre le VHA sont egalement presentes.

20/3,AB/5 (Item 2 from file: 349)

THIS PAGE BLANK (USPTO)

DIALOG(R) File 349:PCT Fulltext
(c) 2001 WIPO/MicroPat. All rts. reserv.

00356498

RECOMBINANT VIRUS IMMUNOTHERAPY
IMMUNOTHERAPIE PAR VIRUS RECOMBINE

Patent Applicant/Assignee:

VIROGENETICS CORPORATION

Inventor(s):

PAOLETTI Enzo

TARTAGLIA James

COX William I

Patent and Priority Information (Country, Number, Date):

Patent: WO 9416716 A1 19940804

Application: WO 94US888 19940121 (PCT/WO US9400888)

Priority Application: US 937115 19930121; US 94184009 19940119

Designated States: AU CA JP AT BE CH DE DK ES FR GB GR IE IT LU MC NL PT SE

Publication Language: English

Fulltext Word Count: 52601

English Abstract

Attenuated recombinant viruses containing DNA coding for a cytokine and/or a tumor associated antigen, as well as methods and compositions employing the viruses, are disclosed and claimed. The recombinant viruses can be NYVAC or ALVAC recombinant viruses. The DNA can code for at least one of: human tumor necrosis factor; nuclear phosphoprotein p53, wildtype or mutant; human melanoma-associated antigen; IL-2; IFNgamma; IL-4; GMCSF; IL-12; B7; erb-B-2 and carcinoembryonic antigen. The recombinant viruses and gene products therefrom are useful for cancer therapy.

Japanese Abstract

Des virus recombinés atténués contenant l'ADN codant pour un antigène associé à une cytokine ou à une tumeur sont décrits et revendiqués, ainsi que des procédés d'utilisation de ces virus et des compositions les comprenant. Les virus recombinés peuvent être des virus NYVAC (virus de la vaccine) ou ALVAC (virus de la variole aviaire). L'ADN peut coder pour au moins l'un des éléments suivants: le facteur de nécrose tumorale humaine, la phosphoprotéine nucléaire p53, de phénotype sauvage ou mutant; l'antigène associé au mélanome humain; l'interleukine-2 (IL-2); l'interféron gamma (IFNgamma); IL-4; le facteur stimulant les colonies de granulocytes macrophages; IL-12; B7; l'antigène carcinoembryonnaire et erb-B-2. Ces virus recombinés et les produits géniques qui en sont dérivés peuvent être utilisés pour la thérapie du cancer.

?

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2001, 16:21:02 ; Search time 121.9 Seconds
(without alignments)
9.379 Million cell updates/sec

Title: US-09-171-432a-39

Perfect score: 107

Sequence: 1 SHTECRPKYKELRLEVQKQR 20

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0401.*

```
1: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	107	100.0	20	Immunogenic Hepati
2	107	100.0	366	Sequence of hepati
3	107	100.0	854	Sequence encoded b
4	107	100.0	993	Sequence of Hepati
5	107	100.0	993	Sequence encoded b
6	107	100.0	1077	A partial hepatitis
7	107	100.0	1091	Translated from 5'
8	107	100.0	2227	Attenuated hepatitis
9	107	100.0	2227	Hepatitis A virus
10	107	100.0	2227	Amino acid sequenc
11	107	100.0	2227	Amino acid sequenc

12	107	100.0	2227	21	B18609	Amino acid sequenc
13	91	85.0	2227	7	P60066	Sequence of viral
14	84.5	79.0	839	12	R15639	Capsid region of c
15	67	62.6	20	18	W42923	Immunogenic Hepati
16	44	41.1	71	21	G07411	Arabidopsis thalia
17	44	41.1	88	21	G61775	Arabidopsis thalia
18	44	41.1	346	12	R11068	1203 antigen seque
19	43	40.2	287	21	G06838	Arabidopsis thalia
20	43	40.2	312	21	G06837	Arabidopsis thalia
21	43	40.2	320	21	G06836	Arabidopsis thalia
22	43	40.2	391	8	P70479	Sequence of the ma
23	43	40.2	391	13	R25305	HRSV major capsid
24	43	40.2	391	19	W50001	HRSV major nucleoc
25	42.5	39.7	5405	16	W14749	IgG-FC binding pro
26	41	38.3	66	21	B42172	Human OREX ORF1936
27	41	38.3	80	21	Y72099	Human serine prote
28	41	38.3	96	17	W03670	HIV-1 strain MAD,
29	41	38.3	112	21	Y91641	Human secreted pro
30	41	38.3	144	21	G08483	Arabidopsis thalia
31	41	38.3	153	21	G08482	Arabidopsis thalia
32	41	38.3	160	20	Y59898	Human normal pancr
33	41	38.3	162	21	G08481	Arabidopsis thalia
34	41	38.3	254	21	B08937	Human secreted pro
35	41	38.3	265	22	B63707	Human gastric canc
36	41	38.3	315	21	B08902	Human secreted pro
37	41	38.3	324	21	B08939	Human secreted pro
38	41	38.3	391	13	R24184	Bovine RSV strain
39	41	38.3	397	13	R21328	Sequence of protei
40	41	38.3	445	21	Y84897	A human proliferat
41	41	38.3	492	20	Y41686	Human PRO274 prote
42	41	38.3	492	21	B44242	Human PRO274 (UNQ2
43	41	38.3	507	21	G40932	Zea mays protein f
44	41	38.3	568	21	G40931	Zea mays protein f
45	41	38.3	568	21	Y94927	Human secreted pro

ALIGNMENTS

```
RESULT 1
W42922
ID W42922 standard; peptide; 20 AA.
XX
AC W42922;
XX
DT 28-APR-1998 (first entry)
XX
DE Immunogenic Hepatitis A virus peptide YK-1315.
XX
KW Immunogenic peptide; immunogenic epitope; P2A protein;
KW Immune response; antibody.
XX
OS Synthetic.
OS Hepatitis A virus.
XX
PN WO9740147-Al.
XX
PD 30-OCT-1997.
XX
PF 18-APR-1997; 97WO-US06891.
XX
PR 19-APR-1996; 96US-0015644.
XX
(PSSH) US DEPT HEALTH & HUMAN SERVICES.
XX
Fields HA, Khudyakov YE;
XX
WPI: 1997-535831/49.
XX
Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an
immune response to HAV in a mammal or to detect the presence of
antibodies against HAV in a mammal
XX
```

PS Claim 18; Page 112; 140pp; English.

XX Peptides W42922-30 are immunogenic peptides corresponding to immunogenic
CC epitopes of the Hepatitis A virus (HAV). The peptides are substantially
CC similar to a portion of the amino acid sequence of the P2A protein of HAV
CC corresponding to amino acids 792-980. The present peptide is derived
CC from amino acids 792-811, and has a reactivity of 54.2% with acute sera.
CC Compositions containing the peptides can be used to induce an immune
CC response to HAV in a mammal. The peptides can also be used to detect the
CC presence of antibodies against HAV in mammalian serum. The peptides can
CC also be used to make an antibody against HAV by administering the peptide
CC to a mammal.

XX Sequence 20 AA;

Query Match 100.0%; Score 107; DB 18; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.4e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHIECRKPYKELRLVGVKQR 20

Db 1 shiecrkpykelrlvvgkqr 20

RESULT 2

P50230
ID P50230 standard; Protein; 366 AA.

XX AC

XX AC

XX DT 28-NOV-1991 (first entry)

XX DE

XX Sequence of hepatitis A virus (HAV) surface protein (VP-1).

XX Hepatitis A virus vaccine; immunisation; monoclonal antibody;

KW diagnostic assay.

XX Hepatitis A virus.

OS Hepatitis A virus.

XX EP138704-A.

XX PD 24-APR-1985.

XX PF 09-OCT-1984; 84EP-0402025.

XX PR 02-MAR-1984; 84US-0585942.

XX PR 14-OCT-1983; 83US-0541836.

XX PA (MERI) MERCK & CO INC.

XX PI Hughes JV, Scolnick EM, Tomassini JE;

XX DR WPI; 1985-100818/17.

XX DR N-PSDB; N50274.

XX PT New hepatitis A virus surface protein - useful for binding to

XX PT neutralising antibodies to the virus

XX PS Claim 21; Page 46-48; 49pp; English.

XX VPI is isolated by solubilisation of the intact virus in an aq.

XX anionic surfactant and a reducing agent. The viral proteins are sepd.

XX CC and the protein of molecular wt. 33000 daltons is sepd.

XX Sequence 366 AA;

Query Match 100.0%; Score 107; DB 6; Length 366;
Best Local Similarity 100.0%; Pred. No. 7.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHIECRKPYKELRLVGVKQR 20

Db 301 shiecrkpykelrlvvgkqr 320

RESULT 3

P50287

ID P50287 standard; Protein; 854 AA.

XX AC

XX AC

XX DT 30-NOV-1991 (first entry)

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX EP154587-A.
 XX 11-SEP-1985.
 XX 27-FEB-1985; 85EP-0400369.
 XX 02-MAR-1984; 84US-0585818.
 XX (MERI) MERCK & CO INC.
 XX Linemeyer DL, Menke JG, Reuben RG, Mitra SW;
 PI WPI: 1985-224964/37.
 DR N-PSDB; N50139.
 XX New nucleotide sequences coding for hepatitis A virus antigens -
 PT useful for eliciting normal immune response and in vaccines for
 PT protecting against the virus
 XX
 PS Example: Page 11-17; 32pp; English.
 XX Within the sequence in N50139 is encoded the information necessary
 CC to make the antigenic proteins of HAV. The sequences encoding for
 CC the structural proteins begin at base 403. The key sub-unit
 CC sequences within VP-1, designated Sequences I,II,III,IV, and V,
 CC start, respectively at 1882, 1963, 1999, 2146, 2347. Other
 CC nucleotide sequences which are valuable as encoding antigenic
 CC proteins are the sequences from base 1749 to base 2722; from base
 CC 1487 to base 2980 and from base 1644 to base 2722. The sequence from
 CC base 1749 to base 2722 is esp. valuable as a vector for producing
 CC antigen protein. Sequences II-V are claimed. X in P50116 denotes the
 CC translation of a stop codon.
 XX Sequence 993 AA;
 SQ

Query Match 100.0%; Score 107; DB 6; Length 993;
 Best Local Similarity 100.0%; Pred. No. 2.1e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVQKQR 20
 ||||||||||||||||
 Db 928 shiecrkpykelrlevgkqr 947

RESULT 5
 P50231
 ID P50231 standard; Protein: 993 AA.
 XX AC
 XX P50231;
 XX 28-NOV-1991 (first entry)
 XX Sequence encoded by partial sequence of hepatitis A virus (HAV),
 DE including surface protein (VP-1).
 XX Hepatitis A virus vaccine; immunisation; monoclonal antibody;
 KW diagnostic assay.
 KW Hepatitis A virus.
 XX Emerson SU, Purcell RH;
 XX WPI: 1999-094412/08.
 FH Key 628..993
 FT Protein /note= "claimed; X denotes translated stop codons
 FT and unspecified triplets"
 XX EP138704-A.
 XX 24-APR-1985.
 XX 09-OCT-1984; 84EP-0402025.
 XX

PR 02-MAR-1984; 84US-0585942.
 PR 14-OCT-1983; 83US-0541836.
 XX (MERI) MERCK & CO INC.
 XX Hughes JV, Scolnick EM, Tomassini JE;
 PI WPI: 1985-100818/17.
 DR N-PSDB; N50274.
 XX New hepatitis A virus surface protein - useful for binding to
 PT neutralising antibodies to the virus
 PT
 XX Disclosure: Page 17-23; 49pp; English.
 XX VP1 is isolated by solubilisation of the intact virus in an aq.
 CC anionic surfactant and a reducing agent. The viral proteins are sepd.
 CC and the protein of molecular wt. 33000 daltons is sepd.
 XX SQ Sequence 993 AA;
 SQ

Query Match 100.0%; Score 107; DB 6; Length 993;
 Best Local Similarity 100.0%; Pred. No. 2.1e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVQKQR 20
 ||||||||||||||||
 Db 928 shiecrkpykelrlevgkqr 947

RESULT 6
 W95559
 ID W95559 standard; Protein: 1077 AA.
 XX AC W95559;
 XX 28-APR-1999 (first entry)
 XX A partial hepatitis A virus (HAV) protein.
 DE Hepatitis A virus protein; HAV; P2 region;
 KW cell-culture-adapted HAV strain; infection; accelerated growth.
 XX Hepatitis A virus.
 OS US5849562-A.
 PN 15-DEC-1998.
 XX 06-JUN-1995; 95US-0468926.
 XX 06-NOV-1991; 91US-0788262.
 PR 30-SEP-1983; 83US-0537911.
 PR 27-SEP-1984; 84US-0654942.
 PR 06-OCT-1988; 88US-0256135.
 PR 06-JUN-1995; 95US-0468926.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Emerson SU, Purcell RH;
 PI WPI: 1999-094412/08.
 DR N-PSDB; X01006.
 XX Chimeric hepatitis A virus strains - with P2 region from
 PT cell-culture-adapted strain in wild-type genome
 PT Disclosure; Fig 7A-L; 36pp; English.
 PS The present sequence represents a partial hepatitis A virus (HAV)
 CC protein. The specification describes a DNA construct consisting
 CC of a wild-type HAV genome in which the P2 region is replaced by the

CC P2 region from a cell-culture-adapted HAV strain. The construct is
 CC used to demonstrate that mutations in the p2 region of a
 CC cell-culture-adapted HAV strain are sufficient for establishment of
 CC infection and accelerated growth in cell culture.

XX SQ Sequence 1077 AA;

Query Match 100.0%; Score 107; DB 20; Length 1077;
 Best Local Similarity 100.0%; Pred. No. 2.2e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLVGVKQR 20
 |||||
 Db 1015 shiecrkpykelrlvkvqr 1034

RESULT 7

ID R32426
 ID R32426 standard; Protein; 1091 AA.

XX AC R32426;

XX DT 10-JUN-1993 (first entry)

XX DE Translated from 5' region of Hepatitis A Virus genomic clone.

XX KW HAV HM-175; chronic liver disease; picornavirus.

XX OS Hepatitis A Virus.

XX FH Key Location/Qualifiers

FT Region 238..1091
 /label= ORF
 /note= "second putative initiation codon at position 240"

FT Region 1..711
 /note= "X's correspond to nonsense codons,
 i.e. this region is not an ORF"

XX US7788262-A.

XX PD 15-DEC-1992.

XX PF 30-SEP-1983; 83US-0536911.

XX PR 27-SEP-1984; 84US-0654942.

XX PR 06-OCT-1988; 88US-0256135.

XX PR 30-SEP-1983; 83US-0536911.

XX PR 06-NOV-1991; 91US-0788262.

XX (USSH) US DEPT HEALTH & HUMAN SERVICE.

XX PI Baltimore D, Feinstein SM;

XX PI Purcell RH, Racaniello VR, Ticehurst JR;

XX DR WPI; 1993-067429/08.

XX DR N-PSDB; Q36934.

XX PT Hepatitis A virus cDNA prodn. - for diagnostic use and for prodn.
 of antigen and antibodies

XX PS Disclosure; Fig 7; 65pp; English.

XX CC HAV virion RNA was extracted from the livers of marmosets which had
 been inoculated with HAV (the HAV had previously been passaged twice
 in marmosets). The RNA was used to prepare ds cDNA clones by
 standard methods. Clones contg. inserts which hybridised to RNA from
 HAV-infected African Green Monkey Kidney cells were selected for
 further analysis. A 7.4kb restriction map (about 99% of the HAV
 genome) was constructed from 5 overlapping inserts. The sequence of
 the first 3.3kb (approx.) from the 5'-terminus was determined. An
 amino acid sequence was deduced from the entire clone and an open

CC reading frame was identified starting at position 238. A comparison
 of the predicted HAV amino acid sequences with the known capsid
 protein sequences of other picornaviruses (poliovirus, foot and
 mouth disease virus and encephalomyelitis virus) revealed areas of
 CC local homology.

XX SQ Sequence 1091 AA;

Query Match 100.0%; Score 107; DB 14; Length 1091;
 Best Local Similarity 100.0%; Pred. No. 2.3e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLVGVKQR 20
 |||||
 Db 1029 shiecrkpykelrlvkvqr 1048

RESULT 8

R05697
 ID R05697 standard; protein; 2227 AA.

XX AC R05697;

XX DT 15-AUG-1990 (first entry)

XX DE Attenuated hepatitis A virus.

XX KW Hepatitis A virus; vaccine; attenuated.

XX OS Hepatitis A virus, strain HM-175.

XX FH Key Location/Qualifiers

FT Region 1..23
 /label=VP4 = 1A

FT Region 24..245
 /label=VP2 = 1B

FT Region 246..491
 /label=VP3 = 1C

FT Region 492..791
 /label=VP1 = 1D

FT Region 792..980
 /label=2A

FT Region 981..1087
 /label=2B

FT Region 1088..1422
 /label=2C

FT Region 1423..1496
 /label=3A

FT Region 1497..1519
 /label=3B = VPg

FT Region 1520..1738
 /label=3C

FT Region 1739..2227
 /label=3D

XX US4894228-A.

XX PD 16-JAN-1990.

XX PF 12-JUL-1988; 88US-0217824.

XX PR 12-JUL-1988; 88US-0217824.

XX PR 12-JUL-1988; 88US-0652967.

XX (USSH) US DEPT HEALTH & HUMAN.

XX PI Purcell RH, Ticehurst JR, Cohen I, Emerson SU, Feinstein SM;

XX PI Daemer RJ, Gust ID;

XX WPI; 1990-075557/10.

XX DR N-PSDB; Q03512.

PT Vaccine against hepatitis A virus infection - comprises novel
 PT attenuated hepatitis A virus strain.
 XX
 PS Claim 1: Fig 1: 18pp; English.
 XX
 CC The attenuated HAV is useful for inducing protective immunity against
 CC HAV. This strain (Pass 35) differs from the wild type HAV HM-175 by
 CC several nucleotide changes distributed throughout the genome, is
 CC attenuated for chimpanzees, elicits serum neutralising antibodies, and is
 CC suitable for use as an HAV vaccine. It is noted that not all the changes
 CC are necessary for attenuation and use as a vaccine.
 XX
 SQ Sequence 2227 AA;

Query Match 100.0%; Score 107; DB 11; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 4.8e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVGKQR 20
 ||||||||||||||||
 Db 792 shiecrkpykelrlevgkqr 811

RESULT 9
 W34074
 ID W34074 standard; Protein: 2227 AA.
 XX
 AC W34074;
 XX
 DT 27-APR-1998 (first entry)
 XX
 XX Hepatitis A virus HM-175 protein sequence.
 DE
 DE HAV; attenuation: 2C protein; 2C chimeric hepatitis A virus;
 KW infection; vaccine.
 KW
 XX Hepatitis A virus HM-175.
 OS

Key Location/Qualifiers
 FH Protein 1..23
 FT /label= VP4
 FT Protein 24..245
 FT /label= VP2
 FT Protein 246..491
 FT /label= VP3
 FT Protein 492..791
 FT /label= VP1
 FT Protein 792..980
 FT /label= 2A
 FT Protein 981..1087
 FT /label= 2B
 FT Protein 1088..1422
 FT /label= 2C
 FT Protein 1423..1496
 FT /label= 3A
 FT Protein 1497..1519
 FT /label= 3B
 FT Protein 1520..1738
 FT /label= 3C
 FT Protein 1739..2227
 FT /label= 3D
 XX WO9740166-A2.
 PN
 XX
 PD 30-OCT-1997.
 XX
 PF 18-APR-1997; 97WO-US06506.
 XX
 PR 19-APR-1996; 96US-0015642.
 XX
 XX (USSH) US SEC DEPT HEALTH.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Emerson SU, Purcell RH, Raychaudhuri G;
 PI WPI; 1997-535850/49.
 XX N-PSDB; T93023.
 DR
 DR Human attenuated HAV genome containing simian HAV 2C gene - useful
 PT as vaccines against HAV infection
 PT
 XX Disclosure; Fig 13A-D; 66pp; English.
 PS
 XX This protein sequence is encoded by the human hepatitis A virus
 CC (HAV) HM-175 wild-type genome (see T93023). Attenuated strain
 CC HAV/7 is obtained by passage of HM-175 in African Green Monkey
 CC kidney cells. A claimed DNA construct (1) comprises a genome of
 CC HAV, where the genome is a human attenuated HAV genome in which a
 CC region of the 2C gene has been replaced by a corresponding region
 CC from a 2C gene of a simian AGM-27 HAV genome (see T93024). The
 CC region of the 2C gene from AGM-27 contained in the construct
 CC preferably encodes amino acids 120-328 of the 2C protein, amino
 CC acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA
 CC transcript of (1); (2) a cell transfected with (1) or the RNA
 CC transcript of (1); (3) a HAV genome as above; (4) antibodies to the
 CC HAV of (3); and (5) a host cell containing the HAV of (3). (1) or
 CC its RNA transcript, can be used as a vaccine for preventing HAV in
 CC a mammal. (1) or the RNA transcript can also be used to stimulate
 CC the production of protective antibodies in the mammal.
 XX
 SQ Sequence 2227 AA;

Query Match 100.0%; Score 107; DB 18; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 4.8e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVGKQR 20
 ||||||||||||||||
 Db 792 shiecrkpykelrlevgkqr 811

RESULT 10
 B18607
 ID B18607 standard; Protein: 2227 AA.
 XX
 AC B18607;
 XX
 DT 15-JAN-2001 (first entry)
 DT
 XX Amino acid sequence of wild type Hepatitis A virus strain HM-175.
 DE
 DE HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.
 KW
 XX Hepatitis A virus.
 OS
 XX US6113912-A.
 PN
 XX 05-SEP-2000.
 PD
 XX 07-JUN-1995; 95US-0475886.
 PF
 XX 18-SEP-1992; 92US-0947338.
 PR
 PR 17-SEP-1993; 93WO-US08610.
 PR
 PR 10-MAR-1995; 95US-0397232.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA
 XX D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
 PI WPI; 2000-586464/55.
 DR N-PSDB; A75476.
 DR
 XX Novel live hepatitis A virus adapted to growth in human fibroblast cell
 PT line useful as vaccine for protecting humans against hepatitis A virus

PT infection, has modified genome compared to wild type -

XX PS Disclosure; Fig 6A-K; 72pp; English.

XX CC The present sequence is derived from a wild type hepatitis A virus
CC (HAV) strain HM-174. The sequence is modified to produce HAV which
CC are adapted to growth in the human fibroblast-like cell line MRC-5.
CC The HAV is able to propagate in MRC-5 cells and retain appropriate
CC attenuation. It is useful as a live vaccine for prophylaxis of
CC hepatitis A in humans and other primates.

XX SQ Sequence 2227 AA;

Query Match 100.0%; Score 107; DB 21; Length 2227;

Best Local Similarity 100.0%; Pred. No. 4.8e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVKGKOR 20

|||||

Db 792 shiecrkpykelrlevgkqr 811

RESULT 11

BI8608

ID BI8608 standard; Protein; 2227 AA.

XX AC

XX BI8608;

DT 15-JAN-2001 (first entry)

XX DE Amino acid sequence of passage 35 Hepatitis A virus called P-35.

XX KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
KW P-35 virus.

XX OS Hepatitis A virus.

XX PN

XX US6113912-A.

XX PD 05-SEP-2000.

XX PF 07-JUN-1995; 95US-0475886.

XX PR 18-SEP-1992; 92US-0947338.

XX PR 17-SEP-1993; 93WO-US08610.

XX PR 10-MAR-1995; 95US-0397232.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;

XX DR WPI; 2000-586464/55.

XX DR N-PSDB; A75477.

XX PT Novel live hepatitis A virus adapted to growth in human fibroblast cell
PT line useful as vaccine for protecting humans against hepatitis A virus
PT infection, has modified genome compared to wild type -

XX PS Disclosure; Columns 67-78; 72pp; English.

XX CC The present sequence is derived from passage 35 of a wild type
CC hepatitis A virus (HAV) strain HM-174. The resulting virus is
CC designated P-35 virus. The sequence is modified to produce HAV which
CC are adapted to growth in the human fibroblast-like cell line MRC-5.
CC The HAV is able to propagate in MRC-5 cells and retain appropriate
CC attenuation. It is useful as a live vaccine for prophylaxis of
CC hepatitis A in humans and other primates.

XX SQ Sequence 2227 AA;

Query Match 100.0%; Score 107; DB 21; Length 2227;

Best Local Similarity 100.0%; Pred. No. 4.8e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVKGKOR 20

|||||

Db 792 shiecrkpykelrlevgkqr 811

RESULT 12

BI8609

ID BI8609 standard; Protein; 2227 AA.

XX AC

XX BI8609;

DT 15-JAN-2001 (first entry)

XX DE Amino acid sequence of live attenuated Hepatitis A virus 4380.

XX KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
KW HAV 4380.

XX OS Hepatitis A virus.

XX PN

XX US6113912-A.

XX PD 05-SEP-2000.

XX PF 07-JUN-1995; 95US-0475886.

XX PR 18-SEP-1992; 92US-0947338.

XX PR 17-SEP-1993; 93WO-US08610.

XX PR 10-MAR-1995; 95US-0397232.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;

XX DR WPI; 2000-586464/55.

XX DR N-PSDB; A75478.

XX PT Novel live hepatitis A virus adapted to growth in human fibroblast cell
PT line useful as vaccine for protecting humans against hepatitis A virus
PT infection, has modified genome compared to wild type -

XX PS Disclosure; Columns 93-104; 72pp; English.

XX CC The present sequence is derived from a live attenuated hepatitis A
CC virus (HAV) of the invention, designated HAV 4380. The sequence is
CC produced by modifying wild type HAV strain HM-174. The HAV of the
CC invention are adapted to growth in the human fibroblast-like cell
CC line MRC-5. The HAV is able to propagate in MRC-5 cells and retain
CC appropriate attenuation. It is useful as a live vaccine for prophylaxis
CC of hepatitis A in humans and other primates.

XX SQ Sequence 2227 AA;

Query Match 100.0%; Score 107; DB 21; Length 2227;

Best Local Similarity 100.0%; Pred. No. 4.8e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVKGKOR 20

|||||

Db 792 shiecrkpykelrlevgkqr 811

RESULT 13

P60066

ID P60066 standard; Protein; 2227 AA.

XX AC

XX P60066;

DT 26-JUN-1991 (first entry)

XX Sequence of viral L434 polypeptide encoded by the complete
DE nucleotide sequence of the HAV genome.
XX Diagnosis: vaccine; passive immunotherapy.
XX Hepatitis A virus.
OS

XX Key Location/Qualifiers
FH Region 1..245
FT /label= P1.1A
FT Region 246..491
FT /label= 1B
FT Region 492..836
FT /label= 1C
FT Region 837..980
FT /label= P2.2A
FT Region 981..1076
FT /label= 2B
FT Region 1077..1422
FT /label= 2C
FT Region 1423..1484
FT /label= P3.3A
FT Region 1485..1507
FT /label= 3B
FT Region 1508..1678
FT /label= 3C
FT Region 1679..2227
FT /label= 3D

XX EPI99480-A.

XX 29-OCT-1986.

XX 03-APR-1986; 86EP-0302465.

XX 03-APR-1985; 85US-0719329.

XX (CHIR-) CHIRON CORP.

XX Dina D, Potter SJ, Vannest GA, Caput D;
XX WPI: 1986-286213/44.

XX N-PSDB; N60080.

XX Hepatitis A virus nucleotide sequence and polypeptide - and use
XX in prodn. of vaccines and diagnostic probes

XX Claim 5; Fig 1; 18pp; English.

XX N60080 and oligonucleotide fragments are useful in detection of
CC hepatitis A virus; transformed hosts may be used for expression of
CC polypeptides and fragments useful in vaccines without risk of
CC infection by the virus or in prodn. of particles which are capable
CC of inducing immunocompetent B cells for passive immunotherapy. Pref.
CC epitope is derived from AAs 445-657 or 792-848 of the HAV
XX polypeptide sequence (P60066).

XX Sequence 2227 AA;

Query Match 85.0%; Score 91; DB 7; Length 2227;
Best Local Similarity 85.0%; Pred. No. 1.5e-05;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLVGKQR 20
Db 792 thiesrkykelrlqvdkqr 811

RESULT 14

RI5629
ID RI5629 standard; Protein; 839 AA.

XX RI5629;
XX 17-MAR-1992 (first entry)
XX Capsid region of cyno-HAV isolate CY-145.
XX Hepatitis A virus; cynomolgus; HAV; monkey; vaccine; macaque.
XX Cynomolgus monkey hepatitis A virus, isolate CY-145.
XX Key Location/Qualifiers
FH Protein 1..245
FT /label= VP0
FT Protein 246..491
FT /label= VP3
FT Protein 492..791
FT /label= VP1
FT Protein 792
FT /label= P2
FT /note= "incomplete"
FT Cleavage-site 245..246
FT Cleavage-site 491..492
FT Cleavage-site 791..792
FT Active-site 315
FT Active-site 593
XX US7678828-A.
XX 12-NOV-1991.
XX 03-APR-1991; 91US-0678828.
XX 03-APR-1991; 91US-0678828.
XX (USSH) US DEPT HEALTH & HUMAN.
XX Hainan OV, Margolis HS, Robertson BH, Brinton MH, Ebert JW;
XX WPI: 1991-376737/51.
XX N-PSDB; Q15180.
XX Hepatitis A virus isolates and DNA - used to prepare vaccines for
XX preventing hepatitis A virus infection.
XX Disclosure; Fig 3; 23pp; English.
XX The sequence was deduced from the nucleotide sequence obt'd. by PCR
XX amplification of cyno-HAV viral RNA obt'd. from the stool of a
XX cynomolgus monkey with serologically and histologically confirmed
XX spontaneous hepatitis A. The sequence differs from the human HAV
XX isolate HM175 (Cohen, J.I., et al. (1987) Proc. Natl. Acad. Sci.
XX USA 84, 2497-2501), mainly in the VP3 and VP1 proteins. The Gln-Val
XX pair at the VP3-VP1 cleavage site in the human isolate is replaced
XX by a Gln-Thr pair in the cyno-HAV. The other two cleavage sites are
XX the same. Two residues have been identified as part of the immuno-
XX dominant region (see feature table) and are different to those in
XX the same position in human HAV. The protein and peptides derived
XX from it can be used in the prepn. of vaccines for the prevention of
XX HAV infection.
XX See also RI5056.
XX Sequence 839 AA;

Query Match 79.0%; Score 84.5; DB 12; Length 839;
Best Local Similarity 90.0%; Pred. No. 5.5e-05;
Matches 18; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 SHIECRKPKYKELRLVGKQR 20
Db 792 shie-kkykelrlvgkqr 810

```

RESULT 15
W42923
ID W42923 standard; peptide; 20 AA.
XX
XX AC W42923;
XX
XX DT 28-APR-1998 (first entry)
XX
XX DE Immunogenic Hepatitis A virus peptide YK-1316.
XX
XX KW Immunogenic peptide; immunogenic epitope; P2A protein;
KW immune response; antibody.
XX
XX OS Synthetic.
XX OS Hepatitis A virus.
XX
XX PN W09740147-A1.
XX
XX PD 30-OCT-1997.
XX
XX PF 18-APR-1997; 97WO-US06891.
XX
XX PR 19-APR-1996; 96US-0015644.
XX
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX PI Fields HA, Khudyakov YE;
XX
XX DR WPI; 1997-535831/49.
XX
XX PT Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an
PT immune response to HAV in a mammal or to detect the presence of
PT antibodies against HAV in a mammal
XX
XX PS Claim 18; Page 112; 140pp; English.
XX
XX CC Peptides W42922-30 are immunogenic peptides corresponding to immunogenic
CC epitopes of the Hepatitis A virus (HAV). The peptides are substantially
CC similar to a portion of the amino acid sequence of the P2A protein of HAV
CC corresponding to amino acids 792-980. The present peptide is derived
CC from amino acids 799-818, and has a reactivity of 41.7% with acute sera.
CC Compositions containing the peptides can be used to induce an immune
CC response to HAV in a mammal. The peptides can also be used to detect the
CC presence of antibodies against HAV in mammalian serum. The peptides can
CC also be used to make an antibody against HAV by administering the peptide
CC to a mammal.
XX
XX SQ Sequence 20 AA;

```

```

Query Match 62.6%; Score 67; DB 18; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00059;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 8 PYKELRLEVGKQR 20
Db 1 pykelrlevgkqr 13
|||||

```

Search completed: April 24, 2001, 16:23:11
Job time: 129 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2001, 16:23:39 ; Search time 113.97 Seconds
(without alignments)
20.568 Million cell updates/sec

Title: US-09-171-432A-39

Perfect score: 107

Sequence: 1 SHIECRKPYKELRLEVKGQR 20

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: SP TREMBL_15.*
- 2: sp_archaea.*
- 3: sp_bacteria.*
- 4: sp_fungi.*
- 5: sp_human.*
- 6: sp_invertebrate.*
- 7: sp_mammal.*
- 8: sp_mhc.*
- 9: sp_organelle.*
- 10: sp_phase.*
- 11: sp_plant.*
- 12: sp_rodent.*
- 13: sp_unclassified.*
- 14: sp_vertebrate.*
- 15: sp_virus.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	100.0	56	14 Q67818	Q67818 hepatitis a
2	107	100.0	56	14 Q67819	Q67819 hepatitis a
3	107	100.0	56	14 Q67820	Q67820 hepatitis a
4	107	100.0	56	14 Q67821	Q67821 hepatitis a
5	107	100.0	56	14 Q67822	Q67822 hepatitis a
6	107	100.0	56	14 Q67823	Q67823 hepatitis a
7	107	100.0	56	14 Q39865	Q39865 hepatitis a
8	107	100.0	56	14 Q39866	Q39866 hepatitis a
9	107	100.0	56	14 Q39867	Q39867 hepatitis a
10	107	100.0	56	14 Q39868	Q39868 hepatitis a
11	107	100.0	56	14 Q39869	Q39869 hepatitis a
12	107	100.0	56	14 Q39870	Q39870 hepatitis a
13	107	100.0	56	14 Q39871	Q39871 hepatitis a
14	107	100.0	56	14 Q39872	Q39872 hepatitis a
15	107	100.0	56	14 Q39873	Q39873 hepatitis a
16	107	100.0	56	14 Q39874	Q39874 hepatitis a
17	107	100.0	56	14 Q39875	Q39875 hepatitis a
18	107	100.0	56	14 Q90P08	Q90P08 hepatitis a
19	107	100.0	56	14 Q91PD2	Q91PD2 hepatitis a

ALIGNMENTS

RESULT 1

Q67818 PRELIMINARY; PRT; 56 AA.
AC Q67818;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
DE POLYPROTEIN (FRAGMENTS).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISR-05, AFULA DISTRICT, ISRAEL, 1993;
RA Karetnyi Y.V., Shulman L.M., Manor J., Leitner L., Shehab S.,
RA Mendelson E.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; 277243; CAB01036.1; -;
DR INTERPRO; IPR000886; -;
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW Polyprotein; Nonstructural protein.
FT NON_TER 1
FT CHAIN 1 >28 CAPSID PROTEIN VP1.
FT NON_CONS 28 29
FT CHAIN 29 >56 NONSTRUCTURAL PROTEIN NS2A.
FT NON_TER 56 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 107; DB 14; Length 56;

Best Local Similarity 100.0%; Pred. No. 1.6e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHIECRKPYKELRLEVKGQR 20

|||||

Db 29 SHIECRKPYKELRLEVKGQR 48

RESULT 2

Q67819 PRELIMINARY; PRT; 56 AA.
ID Q67819

```

AC Q67819;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE POLYPROTEIN (FRAGMENTS).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISR-10, (AFULA DISTRICT, ISRAEL, 1993);
RA Karetnyi Y.V., Shulman L.M., Manor J., Leitner L., Shehab S.,
RA Mendelson E.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; 277244; CAB01037.1; -.
DR INTERPRO; IPR000886; -.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW Polyprotein; Nonstructural protein.
FT NON_TER 1 1
FT CHAIN 1 >28 CAPSID PROTEIN VP1.
FT NON_CONS 28 29
FT CHAIN 29 >56 NONSTRUCTURAL PROTEIN NS2A.
FT NON_TER 56 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 107; DB 14; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECKPKYKELRLEVKGQR 20
Db 29 SHIECKPKYKELRLEVKGQR 48

RESULT 3
Q67820
ID Q67820 PRELIMINARY; PRT; 56 AA.
AC Q67820;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE POLYPROTEIN (FRAGMENTS).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISR-12, (AFULA DISTRICT, ISRAEL, 1993);
RA Karetnyi Y.V., Shulman L.M., Manor J., Leitner L., Shehab S.,
RA Mendelson E.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; 277245; CAB01038.1; -.
DR INTERPRO; IPR000886; -.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW Polyprotein; Nonstructural protein.
FT NON_TER 1 1
FT CHAIN 1 >28 CAPSID PROTEIN VP1.
FT NON_CONS 28 29
FT CHAIN 29 >56 NONSTRUCTURAL PROTEIN NS2A.
FT NON_TER 56 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 107; DB 14; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECKPKYKELRLEVKGQR 20
Db 29 SHIECKPKYKELRLEVKGQR 48

RESULT 4
Q67821
ID Q67821 PRELIMINARY; PRT; 56 AA.
AC Q67821;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE POLYPROTEIN (FRAGMENTS).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISR-15, (AFULA DISTRICT, ISRAEL, 1993);
RA Karetnyi Y.V., Shulman L.M., Manor J., Leitner L., Shehab S.,
RA Mendelson E.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; 277246; CAB01039.1; -.
DR INTERPRO; IPR000886; -.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW Polyprotein; Nonstructural protein.
FT NON_TER 1 1
FT CHAIN 1 >28 CAPSID PROTEIN VP1.
FT NON_CONS 28 29
FT CHAIN 29 >56 NONSTRUCTURAL PROTEIN NS2A.
FT NON_TER 56 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 107; DB 14; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECKPKYKELRLEVKGQR 20
Db 29 SHIECKPKYKELRLEVKGQR 48

RESULT 5
Q67822
ID Q67822 PRELIMINARY; PRT; 56 AA.
AC Q67822;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE POLYPROTEIN (FRAGMENTS).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISR-70, (AKKO DISTRICT, ISRAEL, 1993);
RA Karetnyi Y.V., Shulman L.M., Manor J., Leitner L., Shehab S.,
RA Mendelson E.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; 277247; CAB01040.1; -.
DR INTERPRO; IPR000886; -.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW Polyprotein; Nonstructural protein.
FT NON_TER 1 1
FT CHAIN 1 >28 CAPSID PROTEIN VP1.
FT NON_CONS 28 29
FT CHAIN 29 >56 NONSTRUCTURAL PROTEIN NS2A.
FT NON_TER 56 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 107; DB 14; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECKPKYKELRLEVKGQR 20
Db 29 SHIECKPKYKELRLEVKGQR 48

```



```

FT  NON_TER      1      1
FT  NON_TER     56     56
SQ  SEQUENCE    56 AA;  6614 MW;  8438C51846AEF4A4 CRC64;

Query Match      100.0%; Score 107; DB 14; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 SHIECRKPKYKELRLEVGVKQR 20
    |||||
Db  29 SHIECRKPKYKELRLEVGVKQR 48

RESULT 10
O39868 PRELIMINARY; PRT; 56 AA.
AC O39868;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE POLYPROTEIN (FRAGMENT).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=503712;
RX MEDLINE=97247817; PubMed=9093940;
RA Taylor M.B.;
RT "Molecular epidemiology of South African strains of hepatitis A virus:
RL J. Med. Virol. 51:273-279(1997).
DR EMBL; U68693; AAB53589.1; -.
DR INTERPRO; IPR000886; -.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW Polyprotein.
FT  NON_TER      1      1
FT  NON_TER     56     56
SQ  SEQUENCE    56 AA;  6614 MW;  8438C51846AEF4A4 CRC64;

Query Match      100.0%; Score 107; DB 14; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 SHIECRKPKYKELRLEVGVKQR 20
    |||||
Db  29 SHIECRKPKYKELRLEVGVKQR 48

RESULT 11
O39869 PRELIMINARY; PRT; 56 AA.
AC O39869;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE POLYPROTEIN (FRAGMENT).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=96002382;
RX MEDLINE=97247817; PubMed=9093940;
RA Taylor M.B.;
RT "Molecular epidemiology of South African strains of hepatitis A virus:
RL J. Med. Virol. 51:273-279(1997).
DR EMBL; U68694; AAB53590.1; -.

```

```

DR  INTERPRO; IPR000886; -.
KW  PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
FT  NON_TER      1      1
FT  NON_TER     56     56
SQ  SEQUENCE    56 AA;  6614 MW;  8438C51846AEF4A4 CRC64;

Query Match      100.0%; Score 107; DB 14; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 SHIECRKPKYKELRLEVGVKQR 20
    |||||
Db  29 SHIECRKPKYKELRLEVGVKQR 48

RESULT 12
O39870 PRELIMINARY; PRT; 56 AA.
AC O39870;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE POLYPROTEIN (FRAGMENT).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=406808;
RX MEDLINE=97247817; PubMed=9093940;
RA Taylor M.B.;
RT "Molecular epidemiology of South African strains of hepatitis A virus:
RL J. Med. Virol. 51:273-279(1997).
DR EMBL; U68695; AAB53591.1; -.
DR INTERPRO; IPR000886; -.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW Polyprotein.
FT  NON_TER      1      1
FT  NON_TER     56     56
SQ  SEQUENCE    56 AA;  6642 MW;  465CF51846AEF4BC CRC64;

Query Match      100.0%; Score 107; DB 14; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 SHIECRKPKYKELRLEVGVKQR 20
    |||||
Db  29 SHIECRKPKYKELRLEVGVKQR 48

RESULT 13
O39871 PRELIMINARY; PRT; 56 AA.
AC O39871;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE POLYPROTEIN (FRAGMENT).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=923359;
RX MEDLINE=97247817; PubMed=9093940;
RA Taylor M.B.;
RT "Molecular epidemiology of South African strains of hepatitis A virus:

```



```
RT 1982-1996.";
RL J. Med. Virol. 51:273-279(1997).
DR EMBL: U68696; AAB53592.1; -.
DR INTERPRO: IPR000886; -.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;

Query Match 100.0%; Score 107; DB 14; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVGKQR 20
DB 29 SHIECRKPYKELRLEVGKQR 48
|||||
|||||

RESULT 14
O39872 PRELIMINARY; PRT; 56 AA.
AC O39872;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE POLYPROTEIN (FRAGMENT).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VDM;
RX MEDLINE-97247817; PubMed-9093940;
RA Taylor M.B.;
RT "Molecular epidemiology of South African strains of hepatitis A virus:
1982-1996.";
RL J. Med. Virol. 51:273-279(1997).
DR EMBL: U68697; AAB53593.1; -.
DR INTERPRO: IPR000886; -.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;

Query Match 100.0%; Score 107; DB 14; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVGKQR 20
DB 29 SHIECRKPYKELRLEVGKQR 48
|||||
|||||

RESULT 15
O39873 PRELIMINARY; PRT; 56 AA.
AC O39873;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE POLYPROTEIN (FRAGMENT).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-2333;
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2001, 16:24:44 ; Search time 38.74 Seconds
(without alignments)
17.685 Million cell updates/sec

Title: US-09-171-432a-39

Perfect score: 107

Sequence: 1 SHIECRKPYKELRLEVQKQR 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	107	100.0	341	1 POLG_HPAV1	P13672 hepatitis a
2	107	100.0	852	1 POLG_HPAVC	P06442 hepatitis a
3	107	100.0	2226	1 POLG_HPAV2	P26580 hepatitis a
4	107	100.0	2226	1 POLG_HPAV4	P26581 hepatitis a
5	107	100.0	2226	1 POLG_HPAV8	P26582 hepatitis a
6	107	100.0	2227	1 POLG_HPAVH	P08617 hepatitis a
7	107	100.0	2227	1 POLG_HPAVL	P06441 hepatitis a
8	107	100.0	2227	1 POLG_HPAVM	P13901 hepatitis a
9	95	88.8	808	1 POLG_HPAVG	Q02381 hepatitis a
10	86	80.4	2230	1 POLG_HPAVS	P14553 simian hepa
11	84.5	79.0	839	1 POLG_HPAVT	P31788 simian hepa
12	46	43.0	420	1 YAGA_SCHPO	Q09873 schizosacch
13	43.5	40.7	1663	1 CO3_RAT	P01026 rattus norv
14	43	40.2	391	1 NCAP_HRSV1	P24566 human respi
15	42	39.3	210	1 LOLA_XYLFA	Q9pdc7 xylella fas
16	42	39.3	286	1 PU91_SCICO	P22311 sclara copr
17	42	39.3	723	1 SYM_PYRHO	O58721 pyrococcus
18	42	39.3	861	1 TOPL_BUCAI	P57371 buchnera ap
19	41	38.3	319	1 URED_SYNPV	O87399 synecococc
20	41	38.3	364	1 SUCC_METJA	Q57663 methanococc
21	41	38.3	370	1 T2M2_METJA	Q58844 methanococc
22	41	38.3	391	1 NCAP_BRSV3	P35943 bovine resp
23	41	38.3	391	1 NCAP_BRSV4	P22677 bovine resp
24	41	38.3	391	1 NCAP_BRSVR	O65708 bovine resp
25	41	38.3	391	1 NCAP_HRSVA	P03418 human respi
26	41	38.3	4543	1 LRPI_CHICK	P98157 gallus gall
27	41	38.3	4544	1 LRPI_HUMAN	Q07954 homo sapien
28	40.5	37.9	141	1 V16K_TRVTC	P05076 tobacco rat
29	40.5	37.9	1120	1 RPOM_SCHPO	O13993 schizosacch
30	40.5	37.9	1663	1 CO3_MOUSE	P01027 mus musculu
31	40.5	37.9	2198	1 YLJ2_CAEEL	P34367 caenorhabdi
32	40	37.4	69	1 MOP_HAEIN	P45183 haemophilus
33	40	37.4	309	1 US16_HCMWA	P09717 human cytom

34	40	37.4	345	1 EFB1_MOUSE	P52795 mus musculu
35	40	37.4	391	1 NCAP_ORSVW	Q83957 ovine respi
36	40	37.4	458	1 NIFK_CLOPA	P11347 clostridium
37	40	37.4	558	1 Y369_MVCPN	P75231 mycoplasma
38	40	37.4	1017	1 DPOL_ADEB2	O72539 bovine aden
39	40	37.4	3432	1 POLG_JAEV1	P27395 j genome po
40	40	37.4	3432	1 POLG_JAEV5	P19110 j genome po
41	40	37.4	3432	1 POLG_JAEVJ	P32886 j genome po
42	39.5	36.9	422	1 POLS_RRV2	P7517 ross river
43	39	36.4	141	1 GAT3_YEAST	Q07928 saccharomyc
44	39	36.4	177	1 CYSC_SYNV3	P72940 synecocyst
45	39	36.4	236	1 VBAD_BPP22	Q03546 bacterioph

ALIGNMENTS

```
RESULT 1
POLG_HPAV1
ID POLG_HPAV1 STANDARD: PRT: 341 AA.
AC P13672;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: COAT PROTEINS VP1 TO VP3; CORE PROTEIN
DE P2A] (FRAGMENT).
OS Hepatitis A virus (strain LCDC-1).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus
OX NCBI_TaxID=12093;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89263805; PubMed=2542903;
RA Andonov A.P., Lau P., Chaudhary R.;
RT "Nucleotide sequence of the VP1 gene from a Chinese strain of
RT hepatitis A virus (HAV).";
RL Nucleic Acids Res. 17:3594-3594(1989).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X14666; CAA32794.1; -
CC PIR: S04137; S04137.
CC Polyprotein; Coat protein; Core protein.
KW NON_TER 1 1
FT CHAIN <1 1 COAT PROTEIN VP3 (1C).
FT CHAIN 2 340 COAT PROTEIN VP1 (1D).
FT CHAIN 341 >341 CORE PROTEIN P2A.
FT NON_TER 341 341
SQ SEQUENCE 341 AA; 38003 MW; 066918289BF126D5 CRC64;
```

Query Match 100.0%; Score 107; DB 1; Length 341;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVQKQR 20
|||||
Db 296 SHIECRKPYKELRLEVQKQR 315

RESULT 2
POLG_HPAVC STANDARD: PRT: 852 AA.
ID POLG_HPAVC

```
AC P06442; Q83741; Q83742;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: COAT PROTEINS VP1 TO VP4; CORE PROTEIN
DE P2A] (FRAGMENT).
OS Hepatitis A virus (strain CR326).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12097;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85185648; PubMed=2985793;
RA Linemeyer D.L., Menke J.G., Martin-Gallardo A., Hughes J.V.,
RA Young A., Mitra S.W.;
RT "Molecular cloning and partial sequencing of hepatitis A viral cDNA.";
RL J. Virol. 54:247-255(1985).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M10033; AAA45470.1; -
DR PIR; A03904; GNNYHA.
KW Polyprotein; Coat protein; Core protein.
FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
FT CHAIN 492 836 COAT PROTEIN VP1 (PID).
FT CHAIN 837 >852 CORE PROTEIN P2A.
FT NON_TER 852 852
SQ SEQUENCE 852 AA; 95563 MW; 73D3ED0AD532820E CRC64;

Query Match 100.0%; Score 107; DB 1; Length 852;
Best Local Similarity 100.0%; Pred. No. 1.le-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHIECRKPKYKELRLEVGKOR 20
Db 792 SHIECRKPKYKELRLEVGKOR 811

RESULT 3
POLG_HP2V2
ID POLG_HP2V2 STANDARD; PRT; 2226 AA.
AC P26580;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: COAT PROTEINS VP1 TO VP4; CORE PROTEINS
DE P2A TO P2C; PROBABLE PROTEINS P3A TO P3C; RNA-DIRECTED RNA POLYMERASE
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 24a).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12094;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromeans T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
```

```
recombination.";
RL J. Virol. 65:2056-2065(1991).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS.
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M59810; AAA45468.1; -
DR MEROPS; C03.005; -
DR InterPro; IPR000605; -
DR InterPro; IPR001205; -
DR Pfam; PF006680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_Helicase; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
FT CHAIN 492 794 COAT PROTEIN VP1 (PID).
FT CHAIN 795 900 CORE PROTEIN P2A.
FT CHAIN 901 1087 CORE PROTEIN P2C.
FT CHAIN 1088 1422 PROBABLE PROTEIN P3A.
FT CHAIN 1423 1495 PROBABLE PROTEIN P3B.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3C.
FT CHAIN 1519 1737 RNA-DIRECTED POLYMERASE P3D.
FT CHAIN 1738 2226
SQ SEQUENCE 2226 AA; 251152 MW; 6CD85A91D6B4E2BF CRC64;

Query Match 100.0%; Score 107; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 2.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHIECRKPKYKELRLEVGKOR 20
Db 792 SHIECRKPKYKELRLEVGKOR 811

RESULT 4
POLG_HP4V4
ID POLG_HP4V4 STANDARD; PRT; 2226 AA.
AC P26581;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: COAT PROTEINS VP1 TO VP4; CORE PROTEINS
DE P2A TO P2C; PROBABLE PROTEINS P3A TO P3C; RNA-DIRECTED RNA POLYMERASE
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 43c).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12095;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromeans T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination.";
RL J. Virol. 65:2056-2065(1991).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS.
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
```


DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE GENOME POLYPROTEIN [CONTAINS: COAT PROTEINS VP1 TO VP4; CORE PROTEINS
 DE P2A TO P2C; PROBABLE PROTEINS P3A TO P3C; RNA-DIRECTED RNA POLYMERASE
 DE P3D (EC 2.7.7.48)].
 OS Hepatitis A virus (strain MB8).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12100;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88045071; PubMed=2823500;
 RA Paul A.V., Tada H., der Helm K., Wissel T., Kiehn R., Wimmer E.,
 RA Deinhardt F.;
 RT "The entire nucleotide sequence of the genome of human hepatitis A
 RL virus (isolate MB8).";
 CC VIRUS RES. 8:153-171(1987).
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS.
 CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
 CC VP3, AND VP4.
 CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT
 CC WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M20273; AAA45474.1; -;
 DR PIR; JS0303; GNHYB.
 DR MEROPS; C03.005; -;
 DR InterPro; IPR000605; -;
 DR InterPro; IPR001205; -;
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam; PF00910; RNA_helicase; 1.
 KW Polyprotein; Coat protein; Core protein; Transferase;
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
 FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
 FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
 FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
 FT CHAIN 492 836 COAT PROTEIN VP1 (PID).
 FT CHAIN 837 980 CORE PROTEIN P2A.
 FT CHAIN 981 1087 CORE PROTEIN P2B.
 FT CHAIN 1088 1422 CORE PROTEIN P2C.
 FT CHAIN 1423 1496 PROBABLE PROTEIN P3A.
 FT CHAIN 1497 1519 PROBABLE PROTEIN P3B.
 FT CHAIN 1520 1738 PROBABLE PROTEIN P3C.
 FT CHAIN 1739 2227 RNA-DIRECTED POLYMERASE 3D.
 FT CHAIN 2227 2227 AA; 251425 MW; EC983ED2A7C86349 CRC64;
 SQ SEQUENCE 2227 AA; 251425 MW; EC983ED2A7C86349 CRC64;

Query Match 100.0%; Score 107; DB 1; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 2.8e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVCKQR 20
 DB 792 SHIECRKPYKELRLEVCKQR 811
 RESULT 9
 POLG_HPAVG STANDARD; PRT; 808 AA.
 AC Q02381;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE GENOME POLYPROTEIN [CONTAINS: COAT PROTEINS VP1 TO VP4; CORE PROTEIN
 DE P2A] (FRAGMENT).

OS Hepatitis A virus (strain GA76).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=31706;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92260183; PubMed=1316423;
 RA Khanna B., Spelbring J.E., Innis B.L., Robertson B.H.;
 RT "Characterization of a genetic variant of human hepatitis A virus.";
 RL J. Med. Virol. 36:118-124(1992).
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS.
 CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
 CC VP3, AND VP4.
 CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT
 CC WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M66695; AAA45477.1; -;
 KW Polyprotein; Coat protein; Core protein.
 FT CHAIN 1 1 COAT PROTEIN VP4 (PIA).
 FT CHAIN <1 2 COAT PROTEIN VP2 (PIB).
 FT CHAIN 3 223 COAT PROTEIN VP3 (PIC).
 FT CHAIN 224 470 COAT PROTEIN VP1 (PID).
 FT CHAIN 471 770 COAT PROTEIN VP1 (PID).
 FT CHAIN 771 >808 CORE PROTEIN P2A.
 FT CHAIN 808 808
 SQ SEQUENCE 808 AA; 90632 MW; D80CE7E57A479C12 CRC64;
 Query Match 88.8%; Score 95; DB 1; Length 808;
 Best Local Similarity 95.0%; Pred. No. 9e-08;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVCKQR 20
 DB 771 SHIECRKPYKELRLEVCKQR 790
 RESULT 10
 POLG_HPAVS STANDARD; PRT; 2230 AA.
 AC P14553;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE GENOME POLYPROTEIN [CONTAINS: COAT PROTEINS VP1 TO VP4; CORE PROTEINS
 DE P2A TO P2C; PROBABLE PROTEINS P3A TO P3C; RNA-DIRECTED RNA POLYMERASE
 DE P3D (EC 2.7.7.48)].
 DE P3D (EC 2.7.7.48)].
 OS Simian hepatitis A virus (strain AGM-27).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12102;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91311420; PubMed=1649901;
 RA Tsarev S.A., Emerson S.U., Balayan M.S., Ticehurst J.R.,
 RA Purcell R.H.;
 RT "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome
 RT structure and growth in cell culture with other HAV strains.";
 RL J. Gen. Virol. 72:1677-1683(1991).
 RN [2]
 RP SEQUENCE OF 1750-2164 FROM N.A.
 RX MEDLINE=89232168; PubMed=2541023;
 RA Balayan M.S., Kusov Y.Y., Andjaparidze A.G., Tsarev S.A.,
 RA Sverdlov E.B., Chizhikov V.E., Blinov V.M., Vasilenko S.K.;
 RT "Variations in genome fragments coding for RNA polymerase in human
 RT and simian hepatitis A viruses."

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; D00736; BAA00637.1; -;
 DR PIR; C32063; VHN23.
 KW Nucleocapsid.
 SQ SEQUENCE 391 AA; 43416 MW; 9A2BBB50103B2835 CRC64;

Query Match 40.2%; Score 43; DB 1; Length 391;
 Best Local Similarity 56.2%; Pred. No. 12;
 Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 3 ECRPKYKELRLVGVK 18

||| ||| ||| |||

Db 129 IESRKSRYKLLKEMGE 144

RESULT 15

LOLA_XYLFA

ID LOLA_XYLFA STANDARD; PRT; 210 AA.

AC Q9PDC7;

DT 01-OCT-2000 (Rel. 40, Created)

DT 01-OCT-2000 (Rel. 40, Last sequence update)

DE 01-OCT-2000 (Rel. 40, Last annotation update)

DE OUTER-MEMBRANE LIPOPROTEINS CARRIER PROTEIN PRECURSOR.

GN LOLA OR XF1452.

OS Xylella fastidiosa.

OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;

CC Xylella.

OX NCBI_TaxID=2371;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=9A5C;

RX MEDLINE=20365717; PubMed=10910347;

RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carier H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.P., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohne M., Furlan L.R.,
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hohelsel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.2., Silveira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
 RL Nature 406:151-159(2000).
 CC -!- FUNCTION: PARTICIPATES IN THE TRANSLLOCATION OF LIPOPROTEINS FROM
 CC THE INNER MEMBRANE TO THE OUTER MEMBRANE. ONLY FORMS A COMPLEX
 CC WITH A LIPOPROTEIN IF THE RESIDUE AFTER THE N-TERMINAL CYS IS NOT
 CC AN ASPARTATE (THE ASP ACTS AS A TARGETING SIGNAL TO INDICATE THAT
 CC THE LIPOPROTEIN SHOULD STAY IN THE INNER MEMBRANE) (BY
 CC SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: PERIPLASMIC (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE LOLA FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AE003975; AAF84261.1; ALT_INIT.
 KW Chapterone; Transport; Protein transport; Periplasmic; Signal.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 210 OUTER-MEMBRANE LIPOPROTEINS CARRIER
 FT PROTEIN.
 SQ SEQUENCE 210 AA; 23618 MW; 08EB8BA44005B24F CRC64;

Query Match 39.3%; Score 42; DB 1; Length 210;
 Best Local Similarity 50.0%; Pred. No. 9.3;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 4 ECRPKYKELRLVGVKQ 19

||| ||| ||| |||

Db 73 ECDAPYKQLVVDGKR 88

Search completed: April 24, 2001, 16:33:53
 Job time: 549 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2001, 16:21:06 ; Search time 73.14 Seconds
(without alignments)
18.792 Million cell updates/sec

Title: US-09-171-432A-39
Perfect score: 107
Sequence: 1 SHIECRKPKYKELRLEVQKQR 20
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 67: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	107	100.0	56	2 PQ0431	genome polyprotein
2	107	100.0	341	2 S04137	genome polyprotein
3	107	100.0	852	1 GNNYHA	genome polyprotein
4	107	100.0	1358	2 A03905	genome polyprotein
5	107	100.0	2227	1 GNNYHM	genome polyprotein
6	107	100.0	2227	1 GNNYHR	genome polyprotein
7	107	100.0	2227	1 GNNYMK	genome polyprotein
8	107	100.0	2227	1 GNNYHB	genome polyprotein
9	105	98.1	56	2 PQ0427	genome polyprotein
10	104	97.2	56	2 PQ0429	genome polyprotein
11	102	95.3	319	2 JH0135	genome polyprotein
12	100	93.5	56	2 PQ0430	genome polyprotein
13	97	90.7	56	2 PQ0432	genome polyprotein
14	96	89.7	56	2 PQ0428	genome polyprotein
15	95	88.8	56	2 PQ0434	genome polyprotein
16	87	81.3	55	2 PQ0433	genome polyprotein
17	86	80.4	56	2 PQ0436	genome polyprotein
18	86	80.4	2230	1 GNNYSA	genome polyprotein
19	84.5	79.0	55	2 PQ0435	genome polyprotein
20	84.5	79.0	839	1 GNNYS2	genome polyprotein
21	81	75.7	56	2 PQ0437	genome polyprotein
22	46	43.0	420	2 S62541	hypothetical prote
23	44	41.1	374	2 CB4040	hypothetical prote
24	43.5	40.7	1663	1 C3RT	complement C3 prec
25	43	40.2	339	2 T18926	hypothetical prote
26	43	40.2	346	2 T27896	hypothetical prote
27	43	40.2	391	1 VHNZ1	nucleocapsid prote
28	43	40.2	554	2 T49917	hypothetical prote
29	42.5	39.7	140	2 E64364	hypothetical prote

30	42	39.3	121	2	T17708	hypothetical prote
31	42	39.3	214	2	E82679	chaperone xF1452 I
32	42	39.3	286	2	S07532	puff II/9-1 protel
33	42	39.3	329	2	T25067	hypothetical prote
34	42	39.3	494	2	S62902	legumin 2 precuso
35	42	39.3	723	2	D71091	methionine-tRNA 1
36	42	39.3	1087	2	T16876	hypothetical prote
37	42	39.3	1157	2	T43259	pyruvate (flavodox
38	42	39.3	1270	2	T22615	hypothetical prote
39	41	38.3	298	2	B83823	endonuclease IV BH
40	41	38.3	364	2	C64326	succinyl-CoA synth
41	41	38.3	370	2	H64480	hypothetical prote
42	41	38.3	391	1	VHNZ1	nucleocapsid prote
43	41	38.3	391	1	VHNZB4	nucleocapsid prote
44	41	38.3	391	1	JQ1533	nucleocapsid prote
45	41	38.3	467	1	VHNZ	nucleocapsid prote

ALIGNMENTS

RESULT 1
PQ0431
genome polyprotein - human hepatitis A virus (strain No. 4) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0431
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; W1
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different ge
A:Reference number: PQ0427; MUID:92300330
A:Accession: PQ0431
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A:Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 100.0%; Score 107; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.6e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVQKQR 20
|||||
DB 29 SHIECRKPKYKELRLEVQKQR 48

RESULT 2
S04137
genome polyprotein - human hepatitis A virus (strain LCDC-1) (fragment)
C:Species: human hepatitis A virus
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 21-Jul-2000
C:Accession: S04137
R:Andonov, A.P.; Lau, P.C.K.; Chaudhary, R.
Nucleic Acids Res. 17, 3594, 1989
A:Title: Nucleotide sequence of the VP1 gene from a Chinese strain of hepatitis A vir
A:Reference number: S04137; MUID:89263805
A:Accession: S04137
A:Molecule type: mRNA
A:Residues: 1-341 <AND>
A:Cross-references: EMBL:X14666; NID:g62301; PIDN:CAA32794.1; PID:g4377576
C:Genetics:
A:Gene: VP1
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; polyprotein
F;2-340/Product: coat protein ID (VP1) #status predicted <MAT>

Query Match 100.0%; Score 107; DB 2; Length 341;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHIECRKPKYKELRLEVGKQR 20
|||||
Db 296 SHIECRKPKYKELRLEVGKQR 315

RESULT 3

GNNYHA

genome polyprotein - human hepatitis A virus (strain CR326) (fragment)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Accession: A03904
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999

R:Linemeyer, D.L.; Menke, J.G.; Martin-Gallardo, A.; Hughes, J.V.; Young, A.; Mitra, S.W.
J. Virol. 54, 247-255, 1985
A:Title: Molecular cloning and partial sequencing of hepatitis A viral cDNA.
A:Reference number: A03904; MUID:85185648

A:Accession: A03904

A:Molecule type: genomic RNA

A:Residues: 1-852 <LIN>

A:Cross-references: EMBL:M10033; NID:g329592; PIDN:AAA45470.1; PID:g329593

C:Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; polyprotein

F:1-245/Product: coat protein 1A #status predicted <CIA>

F:246-491/Product: coat protein 1B #status predicted <C1B>

F:492-836/Product: coat protein 1C #status predicted <C1C>

F:837-852/Product: core protein 2A (fragment) #status predicted <C2A>

Query Match 100.0%; Score 107; DB 1; Length 852;

Best Local Similarity 100.0%; Pred. No. 4.2e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHIECRKPKYKELRLEVGKQR 20
|||||
Db 792 SHIECRKPKYKELRLEVGKQR 811

RESULT 4

A03905

genome polyprotein (version 2) - human hepatitis A virus (fragments)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; protein

C:Species: human hepatitis A virus

C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 15-Nov-1996

C:Accession: A03905

R:Baroudy, B.M.; Ticehurst, J.R.; Miele, T.A.; Mairzel Jr., J.V.; Purcell, R.H.; Feinstone

Proc. Natl. Acad. Sci. U.S.A. 82, 2143-2147, 1985

A:Title: Sequence analysis of hepatitis A virus cDNA coding for capsid proteins and RNA

A:Reference number: A03905; MUID:85166289

A:Accession: A03905

A:Molecule type: genomic RNA

A:Residues: 1-1358 <BAR>

C:Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; polyprotein

F:1-245/Product: coat protein 1A #status predicted <CIA>

F:246-491/Product: coat protein 1B #status predicted <C1B>

F:492-836/Product: coat protein 1C #status predicted <C1C>

F:837-854/Product: core protein 2A (fragment) #status predicted <C2A>

F:855-1358/Product: core protein 3D (RNA polymerase) (fragment) #status predicted <C3D>

Query Match

Best Local Similarity 100.0%; Score 107; DB 2; Length 1358;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHIECRKPKYKELRLEVGKQR 20
|||||
Db 792 SHIECRKPKYKELRLEVGKQR 811

RESULT 5

GNNYHM

genome polyprotein - human hepatitis A virus (strain HM-175, wild type)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core
B: RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Accession: A25981
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999

R:Cohen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.

J. Virol. 61, 50-59, 1987

A:Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with

A:Reference number: A25981; MUID:87061253

A:Accession: A25981

A:Molecule type: genomic RNA

A:Residues: 1-2227 <COH>

A:Cross-references: EMBL:M14707; NID:g329582; PIDN:AAA45465.1; PID:g329583

C:Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltr

F:1-23/Product: coat protein 1A #status predicted <VP4>

F:24-245/Product: coat protein 1B #status predicted <VP2>

F:246-491/Product: coat protein 1C #status predicted <VP3>

F:492-791/Product: coat protein 1D #status predicted <VP1>

F:792-980/Product: core protein 2A #status predicted <C2A>

F:981-1087/Product: core protein 2B #status predicted <C2B>

F:1088-1422/Product: core protein 2C #status predicted <C2C>

F:1423-1496/Product: protein 3A #status predicted <C3A>

F:1497-1519/Product: protein 3B #status predicted <C3B>

F:1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>

F:1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 107; DB 1; Length 2227;

Best Local Similarity 100.0%; Pred. No. 1.1e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHIECRKPKYKELRLEVGKQR 20
|||||
Db 792 SHIECRKPKYKELRLEVGKQR 811

RESULT 6

GNNYHR

genome polyprotein - human hepatitis A virus

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core

NA polymerase (EC 2.7.7.48), protein 3D

C:Species: human hepatitis A virus

A:Note: host Homo sapiens (man)

C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999

R:Najarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van N

Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985

A:Title: Primary structure and gene organization of human hepatitis A virus.

A:Reference number: A03903; MUID:85190549

A:Accession: A03903

A:Molecule type: genomic RNA

A:Residues: 1-2227 <NAJ>

A:Cross-references: GB:M02990; NID:g329596; PIDN:AAA45472.1; PID:g329597

C:Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltr

F:1-245/Product: coat protein 1A #status predicted <CIA>

F:246-491/Product: coat protein 1B #status predicted <C1B>

F:492-836/Product: coat protein 1C #status predicted <C1C>

F:837-980/Product: core protein 2A #status predicted <C2A>

F:981-1076/Product: core protein 2B #status predicted <C2B>

F:1077-1422/Product: core protein 2C #status predicted <C2C>

F:1423-1484/Product: core protein 3A #status predicted <C3A>

F:1485-1507/Product: protein 3B #status predicted <C3B>

F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <C3C>

F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match

Best Local Similarity 100.0%; Score 107; DB 1; Length 2227;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGVKQR 20
 |||||
 Db 792 SHIECRKPKYKELRLEVGVKQR 811

RESULT 7
 GNNYMK
 genome polyprotein - human hepatitis A virus (strain HM-175/7MK-5, attenuated HAV)
 N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 2A; core protein 3D
 NA polymerase (EC 2.7.7.48), protein 3D
 C:Species: human hepatitis A virus
 A:Note: host Homo sapiens (man)
 C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999
 C:Accession: A94149; A25914; A94508
 R:Cohen, J.I.; Rosenblum, B.; Ticehurst, J.R.; Daemer, R.J.; Feinstein, S.M.; Purcell, R.
 Proc. Natl. Acad. Sci. U.S.A. 84, 2497-2501, 1987
 A:Title: Complete nucleotide sequence of an attenuated hepatitis A virus: comparison with
 A:Reference number: A94149; MUID:87175701
 A:Accession: A94149
 A:Status: nucleic acid sequence not shown
 A:Molecule type: genomic RNA
 A:Residues: 1-2227 <COH>
 A:Cross-references: EMBL:M16632; NID:g329594; PID:AAA45471.1; PID:g329595
 A:Note: submitted to GenBank, August 1987
 C:Superfamily: hepatitis A virus genome polyprotein
 C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltransferase
 F:1-245/Product: coat protein 1A #status predicted <P1A>
 F:246-491/Product: coat protein 1B #status predicted <P1B>
 F:492-836/Product: coat protein 1C #status predicted <P1C>
 F:837-980/Product: coat protein 2A #status predicted <P2A>
 F:981-1076/Product: core protein 2B #status predicted <P2B>
 F:1077-1422/Product: core protein 2C #status predicted <P2C>
 F:1423-1484/Product: protein 3A #status predicted <P3A>
 F:1485-1507/Product: protein 3B #status predicted <P3B>
 F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <P3C>
 F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 100.0%; Score 107; DB 1; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1.1e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGVKQR 20
 |||||
 Db 792 SHIECRKPKYKELRLEVGVKQR 811

RESULT 8
 GNNYHB
 genome polyprotein - human hepatitis A virus (strain MBB)
 N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein 1E; VPg; protein 3A; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
 C:Species: human hepatitis A virus
 A:Note: host Homo sapiens (man)
 C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-May-1996
 C:Accession: J50303
 R:Paul, A.V.; Tada, H.; von der Helm, K.; Wissel, T.; Kiehn, R.; Wimmer, E.; Deinhardt, R.
 Virus Res. 8, 153-171, 1987
 A:Title: The entire nucleotide sequence of the genome of human hepatitis A virus (isolat
 A:Reference number: JS0303; MUID:88045071
 A:Accession: JS0303
 A:Molecule type: genomic RNA
 A:Residues: 1-2227 <PAU>
 A:Cross-references: EMBL:M20273
 C:Superfamily: hepatitis A virus genome polyprotein
 C:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; hydro
 F:1-23/Product: coat protein 1A #status predicted <VP4>
 F:24-246/Product: coat protein 1B #status predicted <VP2>
 F:247-491/Product: coat protein 1C #status predicted <VP3>
 F:492-836/Product: coat protein 1D #status predicted <VP1>
 F:837-980/Product: core protein 1E #status predicted <P2A>
 F:981-1108/Product: core protein 2B #status predicted <P2B>
 F:1109-1438/Product: core protein 2C #status predicted <P2C>

Query Match 100.0%; Score 107; DB 1; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1.1e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGVKQR 20
 |||||
 Db 792 SHIECRKPKYKELRLEVGVKQR 811

F:1439-1496/Product: protein 3A #status predicted <P3A>
 F:1497-1519/Product: genome-linked protein Vpg #status predicted <VPG>
 F:1520-1736/Product: cysteine proteinase, protein 3C #status predicted <P3C>
 F:1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 100.0%; Score 107; DB 1; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1.1e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGVKQR 20
 |||||
 Db 792 SHIECRKPKYKELRLEVGVKQR 811

RESULT 9
 P00427
 genome polyprotein - human hepatitis A virus (strain EP-35.730) (fragment)
 C:Species: human hepatitis A virus
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
 C:Accession: P00427
 R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; W
 J. Gen. Virol. 73, 1365-1377, 1992
 A:Title: Genetic relatedness of hepatitis A virus strains recovered from different ge
 A:Reference number: P00427; MUID:92300330
 A:Accession: P00427
 A:Molecule type: mRNA
 A:Residues: 1-56 <ROB>
 C:Comment: This protein is from the VP1/2A Junction region.
 C:Superfamily: hepatitis A virus genome polyprotein
 C:Keywords: coat protein; core protein; polyprotein

Query Match 98.1%; Score 105; DB 2; Length 56;
 Best Local Similarity 95.0%; Pred. No. 5.5e-10;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGVKQR 20
 |||||
 Db 29 SHIECRKPKYKELRLEVGVKQR 48

RESULT 10
 P00429
 genome polyprotein - human hepatitis A virus (strain PRC16) (fragment)
 C:Species: human hepatitis A virus
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
 C:Accession: P00429
 R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; W
 J. Gen. Virol. 73, 1365-1377, 1992
 A:Title: Genetic relatedness of hepatitis A virus strains recovered from different ge
 A:Reference number: P00427; MUID:92300330
 A:Accession: P00429
 A:Molecule type: mRNA
 A:Residues: 1-56 <ROB>
 A:Note: This protein is from the VP1/2A Junction region
 C:Superfamily: hepatitis A virus genome polyprotein
 C:Keywords: coat protein; core protein; polyprotein

Query Match 97.2%; Score 104; DB 2; Length 56;
 Best Local Similarity 95.0%; Pred. No. 7.9e-10;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGVKQR 20
 |||||
 Db 29 SHIECRKPKYKELRLEVGVKQR 48

RESULT 11
 JH0135
 genome polyprotein - human hepatitis A virus (strain MS-1) (fragment)
 N:Contains: amino end of core protein 2A; coat protein 1D

C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 08-Apr-1994
C:Accession: JH0135
R:Robertson, B.H.; Brown, V.K.; Bradley, D.W.
Virus Res. 8, 309-316, 1987
A:Title: Nucleic acid sequence of the VP1 region of attenuated MS-1 hepatitis A virus.
A:Reference number: JH0135; MUID:88129044
A:Accession: JH0135
A:Molecule type: genomic RNA
A:Residues: 1-319 <ROB>
A:Cross-references: GB:M22821
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; glycoprotein; polyprotein
F:1-300/Product: coat protein ID #status predicted <CPD>
F:301-319/Product: core protein 2A (fragment) #status predicted <C2A>
F:231/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 95.3%; Score 102; DB 2; Length 319;
Best Local Similarity 100.0%; Pred. No. 9.7e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGKQ 19
||||| ||||||| ||||||| |||
Db 301 SHIECRKPKYKELRLEVGKQ 319

RESULT 12
PQ0430
genome polyprotein - human hepatitis A virus (strain S23-1) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0430
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A:Reference number: PQ0427; MUID:92300330
A:Accession: PQ0430
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A:Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 93.5%; Score 100; DB 2; Length 56;
Best Local Similarity 95.0%; Pred. No. 3.5e-09;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGKQ 20
||||| ||||||| ||||||| |||
Db 29 SHIECGKPKYKELRLEVGKQ 48

RESULT 13
PQ0432
genome polyprotein - human hepatitis A virus (strain CF-53) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0432
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A:Reference number: PQ0427; MUID:92300330
A:Accession: PQ0432
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A:Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 90.7%; Score 97; DB 2; Length 56;
Best Local Similarity 95.0%; Pred. No. 1e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGKQ 20
||||| ||||||| ||||||| |||
Db 29 SHIESRKPYPKELRLEVGKQ 48

RESULT 14
PQ0428
genome polyprotein - human hepatitis A virus (strain TKM002) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0428
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different ge
A:Reference number: PQ0427; MUID:92300330
A:Accession: PQ0428
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A:Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 89.7%; Score 96; DB 2; Length 56;
Best Local Similarity 95.0%; Pred. No. 1.5e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGKQ 20
||||| ||||||| ||||||| |||
Db 29 SHIEYRKPYPKELRLEVGKQ 48

RESULT 15
PQ0434
genome polyprotein - human hepatitis A virus (strain KPH) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0434
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different ge
A:Reference number: PQ0427; MUID:92300330
A:Accession: PQ0434
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A:Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 88.8%; Score 95; DB 2; Length 56;
Best Local Similarity 95.0%; Pred. No. 2.2e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGKQ 20
||||| ||||||| ||||||| |||
Db 29 SHIEKRPYPKELRLEVGKQ 48

Search completed: April 24, 2001, 16:25:42
Job time: 276 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2001, 16:26:09 ; Search time 62.1 Seconds
(without alignments)
6.187 Million cell updates/sec

Title: US-09-171-432A-39

Perfect score: 107
Sequence: 1 SHIECRKPYKELRLEVGKQR 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	100.0	1091	6	Patent No. 5516630
2	107	100.0	2227	4	Sequence 2, Appl
3	107	100.0	2227	4	Sequence 4, Appl
4	107	100.0	2227	4	Sequence 6, Appl
5	107	100.0	2227	4	Sequence 2, Appl
6	107	100.0	2227	4	Sequence 4, Appl
7	84.5	79.0	839	1	Sequence 2, Appl
8	43	40.2	1127	4	Sequence 11, Appl
9	41	38.3	391	5	Sequence 3, Appl
10	41	38.3	4544	1	Sequence 52, Appl
11	41	38.3	4544	2	Sequence 52, Appl
12	39	36.4	24	1	Sequence 13, Appl
13	39	36.4	344	4	Sequence 2, Appl
14	39	36.4	520	6	Sequence 2, Appl
15	39	36.4	1235	2	Patent No. 5223391
16	38	35.5	31	1	Sequence 36, Appl
17	38	35.5	31	1	Sequence 26, Appl
18	38	35.5	102	2	Sequence 15, Appl
19	38	35.5	155	2	Sequence 3, Appl
20	38	35.5	256	2	Sequence 4, Appl
21	38	35.5	256	2	Sequence 2, Appl
22	38	35.5	364	4	Sequence 2, Appl
23	38	35.5	364	4	Sequence 37, Appl
24	38	35.5	372	4	Sequence 24, Appl
25	37.5	35.0	309	4	Sequence 2, Appl
26	37	34.6	67	6	Patent No. 5196333
27	37	34.6	68	1	Sequence 8, Appl

28	37	34.6	68	2	US-09-111-348-8	Sequence 8, Appl
29	37	34.6	69	5	PCT-US95-06406A-6	Sequence 6, Appl
30	37	34.6	469	2	US-08-968-751-2	Sequence 2, Appl
31	37	34.6	493	6	5196333-4	Patent No. 5196333
32	37	34.6	855	2	US-08-468-558-2	Sequence 2, Appl
33	37	34.6	972	3	US-08-335-844A-24	Sequence 24, Appl
34	36.5	34.1	188	4	US-08-933-750C-8	Sequence 8, Appl
35	36.5	34.1	188	4	US-09-234-613-8	Sequence 8, Appl
36	36.5	34.1	609	2	US-08-716-301-4	Sequence 4, Appl
37	36	33.6	102	3	US-08-906-769-125	Sequence 125, App
38	36	33.6	102	4	US-08-906-616-125	Sequence 125, App
39	36	33.6	102	4	US-08-639-075A-125	Sequence 125, App
40	36	33.6	102	4	US-09-012-431-125	Sequence 125, App
41	36	33.6	183	1	US-07-841-646-23	Sequence 23, Appl
42	36	33.6	183	1	US-08-147-023-23	Sequence 23, Appl
43	36	33.6	183	1	US-08-447-570-23	Sequence 23, Appl
44	36	33.6	183	2	US-08-449-700-23	Sequence 23, Appl
45	36	33.6	183	2	US-08-449-699A-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1
5516630-2
; Patent No. 5516630
; APPLICANT: TICEHURST, JOHN R.; BALTIMORE, DAVID; FEINSTONE, STEPHEN M.; PURCELL, ROBERT H.; RACANELLO, VINCENT R.; BAROUDY, BAHIGE M.
; TITLE OF INVENTION: METHODS OF DETECTING HEPATITIS A VIRUS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/788,262
; FILING DATE: 06-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 256,135
; FILING DATE: 06-OCT-1988
; APPLICATION NUMBER: 654,942
; FILING DATE: 27-SEP-1984
; APPLICATION NUMBER: 537,911
; FILING DATE: 30-SEP-1983
; SEQ ID NO: 2:
; LENGTH: 1091
5516630-2

Query Match 100.0%; Score 107; DB 6; Length 1091;
Best Local Similarity 100.0%; Pred. No. 4e-10; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

QY 1 SHIECRKPYKELRLEVGKQR 20
|||||
Db 1029 SHIECRKPYKELRLEVGKQR 1048

RESULT 2
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2

Query Match      100.0%; Score 107; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8.7e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHIECRKPKYKELRLEVKGQR 20
    |||||||||||||||||||
Db 792 SHIECRKPKYKELRLEVKGQR 811

RESULT 3
US-08-475-886-4
; Sequence 4, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475.886A
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947.338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397.232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4

Query Match      100.0%; Score 107; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8.7e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHIECRKPKYKELRLEVKGQR 20
    |||||||||||||||||||
Db 792 SHIECRKPKYKELRLEVKGQR 811

RESULT 4
US-08-475-886-6
; Sequence 6, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475.886A
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947.338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397.232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6

; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6

Query Match      100.0%; Score 107; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8.7e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHIECRKPKYKELRLEVKGQR 20
    |||||||||||||||||||
Db 792 SHIECRKPKYKELRLEVKGQR 811

RESULT 5
US-08-397-232-2
; Sequence 2, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397.232A
; EARLIER FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947.338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2

Query Match      100.0%; Score 107; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8.7e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHIECRKPKYKELRLEVKGQR 20
    |||||||||||||||||||
Db 792 SHIECRKPKYKELRLEVKGQR 811

RESULT 6
US-08-397-232-4
; Sequence 4, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397.232A
; EARLIER FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947.338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4
```


Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 3 ICRPKYKELRLEVCK 18
|| || ||: ||:
Db 129 IESRSYKMLKEMGE 144

RESULT 10

US-08-469-486-52
; Sequence 52, Application US/08469486
; Patent No. 5739281

; GENERAL INFORMATION:

; APPLICANT: Thoeerssen, Hans Christian

; APPLICANT: Holtet, Thor Las

; APPLICANT: Etzerodt, Michael

; TITLE OF INVENTION: Improved method for the refolding of

; TITLE OF INVENTION: proteins

; NUMBER OF SEQUENCES: 58

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version

; SOFTWARE: #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/469,486

; FILING DATE:

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/192,060

; FILING DATE: February 4, 1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Paul T. Clark

; REGISTRATION NUMBER: 30,162

; REFERENCE/DOCKET NUMBER: 06363/002001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617 542 5070

; TELEFAX: 617 542 8906

; TELEX: 200154

; INFORMATION FOR SEQ ID NO: 52:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4544 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-469-486-52

Query Match 38.3%; Score 41; DB 1; Length 4544;

Best Local Similarity 56.2%; Pred. No. 3.6e+02;

Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 5 CRPKYKELRLEVCK 20

|| || || || ||

Db 519 CKKPEHFLVYGKGR 534

RESULT 11

US-08-469-658-52

; Sequence 52, Application US/08469658

; Patent No. 5917018

; GENERAL INFORMATION:

; APPLICANT: Thoeerssen, Hans Christian

; APPLICANT: Holtet, Thor Las

; APPLICANT: Etzerodt, Michael

; TITLE OF INVENTION: IMPROVED METHOD FOR THE REFOOLDING OF

; TITLE OF INVENTION: PROTEINS

; NUMBER OF SEQUENCES: 58

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version

; SOFTWARE: #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/469,658

; FILING DATE: June 5, 1995

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/192,060

; FILING DATE: February 4, 1994

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: Paul T. Clark

; REGISTRATION NUMBER: 30,162

; REFERENCE/DOCKET NUMBER: 06363/002002

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617 542 5070

; TELEFAX: 617 542 8906

; TELEX: 200154

; INFORMATION FOR SEQ ID NO: 52:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4544 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-469-658-52

Query Match

Best Local Similarity 56.2%; Score 41; DB 2; Length 4544;

Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 5 CRPKYKELRLEVCK 20

|| || || || ||

Db 519 CKKPEHFLVYGKGR 534

RESULT 12

US-08-406-347A-13

; Sequence 13, Application US/08406347A

; Patent No. 5714374

; GENERAL INFORMATION:

; APPLICANT: ARNOLD, Edward V.

; APPLICANT: ARNOLD, Gall F.

; TITLE OF INVENTION: CHIMERIC RHINOVIRUSES

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Kenyon & Kenyon

; STREET: 1025 Connecticut Avenue, N.W., Suite 600

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.50 inch floppy disk, 1.44 mb storage

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Wordperfect 6.1 Windows

; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/406,347A
FILING DATE: 17-MAR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/304,635
APPLICATION DATE: 08/041,790
APPLICATION NUMBER: 08/041,790
APPLICATION DATA: 07/583,335
ATTORNEY/AGENT INFORMATION:
NAME: Toffenetti, Judith L.
REGISTRATION NUMBER: 39,048
REFERENCE/DOCKET NUMBER: 1984/46203
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-429-1776
TELEFAX: 202-429-0796
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-406-347A-13

Query Match 36.4%; Score 39; DB 1; Length 24;
Best Local Similarity 33.3%; Pred. No. 2.5;
Matches 6; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 3 IECKPKYKELRLEVGVKQR 20
DB 5 INCTPDYNNVRSLSRTR 22

RESULT 13
US-09-393-554-2
Sequence 2, Application US/09393554
Patent No. 6210897
GENERAL INFORMATION:
APPLICANT: Andersson, Leif
APPLICANT: Kijas, James
APPLICANT: Gafvert, Sophie
APPLICANT: Wigh-Trowaldh, Gunilla
APPLICANT: Hedhammar, Ake
TITLE OF INVENTION: IDENTIFICATION OF CANINE LEUKOCYTE ADHESION DEFICIENCY
FILE REFERENCE: 201515/1001
CURRENT APPLICATION NUMBER: US/09/393,554
CURRENT FILING DATE: 1999-09-10
EARLIER APPLICATION NUMBER: 60/136,099
EARLIER FILING DATE: 1999-05-26
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 344
TYPE: PRT
ORGANISM: Bos sp.
US-09-393-554-2

Query Match 36.4%; Score 39; DB 4; Length 344;
Best Local Similarity 32.1%; Pred. No. 46;
Matches 9; Conservative 3; Mismatches 4; Indels 12; Gaps 1;

QY 4 ECKPKY-----KELRLEVGVKQ 19
DB 197 ECQPPAFRHLVRLKLTNSKQFETEVGVKQ 224

RESULT 14
5223391-5
Patent No. 5223391
APPLICANT: COEN, DONALD M.; DIGARD, PAUL E.

TITLE OF INVENTION: INHIBITORS OF HERPES SIMPLEX VIRUS
REPLICATION
NUMBER OF SEQUENCES: 9
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/482,634
FILING DATE: 21-FEB-1990
SEQ ID NO: 5
LENGTH: 520
5223391-5

Query Match 36.4%; Score 39; DB 6; Length 520;
Best Local Similarity 45.0%; Pred. No. 73;
Matches 9; Conservative 4; Mismatches 5; Indels 2; Gaps 1;

QY 1 SHIECKPKYKELRLEVGVKQR 20
DB 295 AHLEAGKDY--LEIEVGGRR 312

RESULT 15
US-08-680-326-36
Sequence 36, Application US/08680326
Patent No. 5925733
GENERAL INFORMATION:
APPLICANT: ROSE, TIMOTHY M.
APPLICANT: BOSCH, MARNIX
APPLICANT: STRAND, KURT
APPLICANT: TODARO, GEORGE J.
TITLE OF INVENTION: DNA POLYMERASE OF GAMMA HERPES VIRUSES
TITLE OF INVENTION: ASSOCIATED WITH KAPOSI'S SARCOMA AND RETROPERITONEAL
TITLE OF INVENTION: FIBROMATOSIS
NUMBER OF SEQUENCES: 152
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/680,326
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Schiff, J. Michael
REGISTRATION NUMBER: 40,253
REFERENCE/DOCKET NUMBER: 29938-20001.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 1235 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-680-326-36

Query Match 36.4%; Score 39; DB 2; Length 1235;
Best Local Similarity 45.0%; Pred. No. 1.9e+02;
Matches 9; Conservative 4; Mismatches 5; Indels 2; Gaps 1;

QY 1 SHIECKPKYKELRLEVGVKQR 20
DB 744 AHLEAGKDY--LEIEVGGRR 761

Search completed: April 24, 2001, 16:34:58
Job time: 529 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2001, 16:23:11 ; Search time 121.9 Seconds
(without alignments)
9.379 Million cell updates/sec

Title: US-09-171-432A-40

Perfect score: 102

Sequence: 1 PYKELRLEVGKRLKYAQEE 20

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_0401.*

1: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	102	100.0	20	18 W42923	Immunogenic Hepati
2	102	100.0	366	6 P50230	Sequence of hepati
3	102	100.0	854	6 P50287	Sequence encoded b
4	102	100.0	993	6 P50116	Sequence of Hepati
5	102	100.0	993	6 P50231	Sequence encoded b
6	102	100.0	1077	20 W95559	A partial hepatiti
7	102	100.0	1091	14 R32426	Translated from 5'
8	102	100.0	2227	11 R32426	Attenuated hepati
9	102	100.0	2227	11 R32426	Hepatitis A virus
10	102	100.0	2227	21 B18607	Amino acid sequenc
11	102	100.0	2227	21 B18608	Amino acid sequenc

12	102	100.0	2227	21	B18609	Amino acid sequenc
13	99	97.1	2227	7	P60066	Sequence of viral
14	94	92.2	839	12	R15629	Capsid region of c
15	67	65.7	20	18	W42922	Immunogenic Hepati
16	45	44.1	20	18	W42924	Immunogenic Hepati
17	45	44.1	25	18	W42930	Immunogenic Hepati
18	45	44.1	124	20	Y11920	Human 5' EST seque
19	45	44.1	214	20	Y34690	Chlamydia pneumoni
20	42	41.2	411	21	G39793	Human 5' EST seque
21	42	41.2	452	21	G39792	Arabidopsis thalia
22	41	40.2	68	18	W09415	Arabidopsis thalia
23	41	40.2	68	18	W09417	Human G protein ga
24	41	40.2	68	19	W61906	Human G protein ga
25	41	40.2	68	20	Y17986	Human G protein ga
26	41	40.2	68	20	Y14069	Human CBLAEH07 pro
27	41	40.2	79	21	B43432	Human cancer assoc
28	41	40.2	88	21	G00169	Human secreted pro
29	41	40.2	147	17	R88419	Human 5 lipoxigena
30	41	40.2	161	21	B33220	Eucalyptus grandis
31	41	40.2	187	21	B53341	Human colon cancer
32	40.5	39.7	795	21	Y70962	Rat Ras signalling
33	40	39.2	189	20	Y18062	S. pneumoniae pth
34	40	39.2	189	20	Y18063	S. pneumoniae pth
35	40	39.2	203	21	B51986	Human secreted pro
36	40	39.2	272	21	B54209	Human pancreatic c
37	40	39.2	287	21	Y95447	Putative human des
38	40	39.2	306	19	W84153	Human desaturase e
39	40	39.2	306	20	W95511	Amino acid sequenc
40	40	39.2	306	20	W85132	A desaturase enzym
41	40	39.2	306	21	Y84700	Amino acid sequenc
42	40	39.2	306	21	Y92615	Human desaturase h
43	40	39.2	328	21	Y95449	Human delta-5-desa
44	40	39.2	444	21	B41998	Human ORFX ORF1762
45	40	39.2	444	21	Y97538	Human fatty acid d

ALIGNMENTS

RESULT 1
ID W42923 standard; peptide; 20 AA.
XX
AC W42923;
XX
DT 28-APR-1998 (first entry)
XX
DE Immunogenic Hepatitis A virus peptide YK-1316.
XX
KW Immunogenic peptide; immunogenic epitope; P2A protein;
KW Immune response; antibody.
XX
OS Synthetic.
OS Hepatitis A virus.
XX
PN W09740147-A1.
XX
PD 30-OCT-1997.
XX
PF 18-APR-1997; 97WO-US06891.
XX
PR 19-APR-1996; 96US-0015644.
XX
(USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Fields HA, Khudyakov YE;
XX
DR WPI; 1997-535831/49.
XX
PT Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an
PT Immune response to HAV in a mammal or to detect the presence of
XX antibodies against HAV in a mammal
XX

PS Claim 18; Page 112; 140pp; English.

XX Peptides W42922-30 are immunogenic peptides corresponding to immunogenic
CC epitopes of the Hepatitis A virus (HAV). The peptides are substantially
CC similar to a portion of the amino acid sequence of the P2A protein of HAV
CC corresponding to amino acids 792-980. The present peptide is derived
CC from amino acids 799-818, and has a reactivity of 41.7% with acute sera.
CC Compositions containing the peptides can be used to induce an immune
CC response to HAV in a mammal. The peptides can also be used to detect the
CC presence of antibodies against HAV in mammalian serum. The peptides can
CC also be used to make an antibody against HAV by administering the peptide
CC to a mammal.

XX Sequence 20 AA;

Query Match 100.0%; Score 102; DB 18; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.4e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVKGKRLKYAQEE 20
DB 1 pykelrlevgkqrlkyaqee 20
|||||

RESULT 2
P50230
ID P50230 standard; Protein; 366 AA.

XX AC P50230;

XX 28-NOV-1991 (first entry)

XX Sequence of hepatitis A virus (HAV) surface protein (VP-1).

XX Hepatitis A virus vaccine; immunisation; monoclonal antibody;
KW diagnostic assay.

XX Hepatitis A virus.

XX EPI38704-A.

XX 24-APR-1985.

XX 09-OCT-1984; 84EP-0402025.

XX 02-MAR-1984; 84US-0585942.

XX 14-OCT-1983; 83US-0541836.

XX (MERI) MERCK & CO INC.

XX Hughes JV, Scolnick EM, Tomassini JE;

XX WPI; 1985-100818/17.

XX N-PSDB; N50274.

XX New hepatitis A virus surface protein - useful for binding to
PT neutralising antibodies to the virus

XX Claim 21; Page 46-48; 49pp; English.

XX VP1 is isolated by solubilisation of the intact virus in an aq.
CC anionic surfactant and a reducing agent. The viral proteins are sepd.
CC and the protein of molecular wt. 33000 daltons is sepd.

XX Sequence 366 AA;

Query Match 100.0%; Score 102; DB 6; Length 366;
Best Local Similarity 100.0%; Pred. No. 9.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVKGKRLKYAQEE 20

DB 308 pykelrlevgkqrlkyaqee 327
|||||

RESULT 3

ID P50287
P50287 standard; Protein; 854 AA.

XX AC P50287;

XX 30-NOV-1991 (first entry)

XX Sequence encoded by hepatitis A virus (HAV) cDNA from near the
DE genome 5' terminus to the end of the area corresponding to the
DE capsid protein region of poliovirus RNA.

XX Hepatitis A virus assay; antigen; antibody.

XX Hepatitis A virus.

XX W08501517-A.

XX 11-APR-1985.

XX 27-SEP-1984; 84WO-US01552.

XX 30-SEP-1983; 83US-0537911.

XX (MASI) MASSACHUSETTS INST TECH.

XX Ticehurst JR, Baltimore D, Feinstein SM, Purcell RH;

XX Racaniello VR;

XX WPI; 1985-098846/16.

XX N-PSDB; N50330.

XX New hepatitis A virus CDNA - useful in assays for the virus and
PT for prodn. of the viral antigen and antibodies to it

XX Example; Fig 7; 60pp; English.

XX The inventors claim HAV cDNA and a method for producing it, whereby
CC large amts. can be obtd. economically. The cDNA is useful in the
CC assay for detection of HAV quickly and easily and with high
CC sensitivity and specificity. The HAV cDNA is also used in the prodn.
CC of HAV antigen or antibodies to it. The antibodies may be monoclonal.

XX Sequence 854 AA;

Query Match 100.0%; Score 102; DB 6; Length 854;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVKGKRLKYAQEE 20
|||||

DB 799 pykelrlevgkqrlkyaqee 818
|||||

RESULT 4

ID P50116
P50116 standard; Protein; 993 AA.

XX AC P50116;

XX 30-SEP-1991 (first entry)

XX Sequence of Hepatitis A virus (HAV) immunogenic peptides
DE VP-1, VP-2, VP-3 and VP-4.

XX Antigenic protein; immunogen; vaccine.

XX Hepatitis A virus (strain CR326).

XX EP154587-A.
 XX
 PD 11-SEP-1985.
 XX
 XX 27-FEB-1985; 85EP-0400369.
 XX
 PR 02-MAR-1984; 84US-0585818.
 XX
 XX (MERI) MERCK & CO INC.
 XX
 XX Linemeyer DL, Menke JG, Reuben RG, Mitra SW;
 PI
 XX WPI; 1985-224964/37.
 DR N-PSDB; N50139.
 XX
 XX New nucleotide sequences coding for hepatitis A virus antigens -
 PT useful for eliciting normal immune response and in vaccines for
 PT protecting against the virus
 XX
 PS Example; Page 11-17; 32pp; English.
 PS
 CC Within the sequence in N50139 is encoded the information necessary
 CC to make the antigenic proteins of HAV. The sequences encoding for
 CC the structural proteins begin at base 403. The key sub-unit
 CC sequences within VP-1, designated Sequences I, II, III, IV, and V,
 CC start, respectively at 1882, 1963, 1999, 2146, 2347. Other
 CC nucleotide sequences which are valuable as encoding antigenic
 CC proteins are the sequences from base 1749 to base 2722; from base
 CC 1487 to base 2980 and from base 1644 to base 2722. The sequence from
 CC base 1749 to base 2722 is esp. valuable as a vector for producing
 CC antigen protein. Sequences II-V are claimed. X in P50116 denotes the
 CC translation of a stop codon.
 XX
 XX Sequence 993 AA;
 SQ

Query Match 100.0%; Score 102; DB 6; Length 993;
 Best Local Similarity 100.0%; Pred. No. 2.8e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
 |||||
 Db 935 pykelrlevgkqrlyaqee 954

RESULT 5
 P50231
 ID P50231 standard; Protein: 993 AA.
 XX
 AC P50231;
 XX
 XX 28-NOV-1991 (first entry)
 DT
 XX
 DE Sequence encoded by partial sequence of hepatitis A virus (HAV),
 DE including surface protein (VP-1).
 XX
 XX Hepatitis A virus vaccine; immunisation; monoclonal antibody;
 KW diagnostic assay.
 KW
 XX Hepatitis A virus.
 OS
 XX Key Location/Qualifiers
 FH Protein 628..993
 FT /note= "claimed; X denotes translated stop codons
 FT and unspecified triplets"
 XX
 XX EP138704-A.
 PN
 XX 24-APR-1985.
 PD
 XX
 XX 09-OCT-1984; 84EP-0402025.
 XX
 XX

PR 02-MAR-1984; 84US-0585942.
 XX 14-OCT-1983; 83US-0541836.
 XX
 PA (MERI) MERCK & CO INC.
 XX
 XX Hughes JV, Scolnick EM, Tomassini JE;
 PI
 XX WPI; 1985-100818/17.
 DR N-PSDB; N50274.
 XX
 XX New hepatitis A virus surface protein - useful for binding to
 PT neutralising antibodies to the virus
 PT
 XX Disclosure; Page 17-23; 49pp; English.
 PS
 XX VP1 is isolated by solubilisation of the intact virus in an aq.
 CC anionic surfactant and a reducing agent. The viral proteins are sepd.
 CC and the protein of molecular wt. 33000 daltons is sepd.
 XX
 XX Sequence 993 AA;
 SQ

Query Match 100.0%; Score 102; DB 6; Length 993;
 Best Local Similarity 100.0%; Pred. No. 2.8e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
 |||||
 Db 935 pykelrlevgkqrlyaqee 954

RESULT 6
 W95559
 ID W95559 standard; Protein: 1077 AA.
 XX
 AC W95559;
 XX
 XX 28-APR-1999 (first entry)
 DT
 XX
 DE A partial hepatitis A virus (HAV) protein.
 XX
 KW Hepatitis A virus protein; HAV; P2 region;
 KW cell-culture-adapted HAV strain; infection; accelerated growth.
 XX
 OS Hepatitis A virus.
 XX
 XX US5849562-A.
 PN
 XX 15-DEC-1998.
 PD
 XX
 XX 06-JUN-1995; 95US-0468926.
 PF
 XX
 XX 06-NOV-1991; 91US-0788262.
 PR
 XX 30-SEP-1983; 83US-0537911.
 PR
 XX 27-SEP-1984; 84US-0654942.
 PR
 XX 06-OCT-1988; 88US-0256135.
 PR
 XX 06-JUN-1995; 95US-0468926.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA
 XX Emerson SU, Purcell RH;
 PI
 XX WPI; 1999-094412/08.
 DR N-PSDB; X01006.
 XX
 XX Chimeric hepatitis A virus strains - with P2 region from
 PT cell-culture-adapted strain in wild-type genome
 PT
 XX Disclosure; Fig 7A-L; 36pp; English.
 PS
 XX The present sequence represents a partial hepatitis A virus (HAV)
 CC protein. The specification describes a DNA construct consisting
 CC of a wild-type HAV genome in which the P2 region is replaced by the

CC P2 region from a cell-culture-adapted HAV strain. The construct is
 CC used to demonstrate that mutations in the P2 region of a
 CC cell-culture-adapted HAV strain are sufficient for establishment of
 CC infection and accelerated growth in cell culture.

XX SQ Sequence 1077 AA;

Query Match 100.0%; Score 102; DB 20; Length 1077;
 Best Local Similarity 100.0%; Pred. No. 3e-08; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 0;

Oy 1 PYKELRLVGVKQRLKYAQEE 20
 |||||
 Db 1022 pykelrlvkvqkrikyaqee 1041

RESULT 7

ID R32426 standard; Protein; 1091 AA.

XX R32426;

DT 10-JUN-1993 (first entry)

DE Translated from 5' region of Hepatitis A Virus genomic clone.

KW HAV HM-175; chronic liver disease; picornavirus.

XX Hepatitis A Virus.

XX Key Location/Qualifiers

FT Region 238..1091

FT /label= ORF

FT /note= "second putative initiation codon at position 240"

FT Region 1..711

FT /note= "X's correspond to nonsense codons, i.e. this region is not an ORF"

XX US7788262-A.

XX 15-DEC-1992.

XX 30-SEP-1983; 83US-0536911.

XX 27-SEP-1984; 84US-0654942.

XX 06-OCT-1988; 88US-0256135.

XX 30-SEP-1983; 83US-0536911.

XX 06-NOV-1991; 91US-0788262.

XX (USSH) US DEPT HEALTH & HUMAN SERVICE.

XX Baltimore D, Feinstein SM;

PI Purcell RH, Racaniello VR, Ticehurst JR;

XX WPI: 1993-067429/08.

XX N-PSDB; Q36934.

PT Hepatitis A virus cDNA prodn. - for diagnostic use and for prodn.
 of antigen and antibodies

XX Disclosure; Fig 7; 65pp; English.

XX HAV virion RNA was extracted from the livers of marmosets which had
 CC been inoculated with HAV (the HAV had previously been passaged twice
 CC in marmosets). The RNA was used to prepare ds cDNA clones by
 CC standard methods. Clones contg. inserts which hybridised to RNA from
 CC HAV-infected African Green Monkey kidney cells were selected for
 CC further analysis. A 7.4kb restriction map (about 99% of the HAV
 CC genome) was constructed from 5 overlapping inserts. The sequence of
 CC the first 3.3kb (approx.) from the 5'-terminus was determined. An
 CC amino acid sequence was decoded from the entire clone and an open

CC reading frame was identified starting at position 238. A comparison
 CC of the predicted HAV amino acid sequences with the known capsid
 CC protein sequences of other picornaviruses (poliovirus, foot and
 CC mouth disease virus and encephalomyelitis virus) revealed areas of
 CC local homology.

XX SQ Sequence 1091 AA;

Query Match 100.0%; Score 102; DB 14; Length 1091;
 Best Local Similarity 100.0%; Pred. No. 3.1e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PYKELRLVGVKQRLKYAQEE 20
 |||||
 Db 1036 pykelrlvkvqkrikyaqee 1055

RESULT 8

ID R05697 standard; protein; 2227 AA.

XX R05697;

DT 15-AUG-1990 (first entry)

DE Attenuated hepatitis A virus.

XX Hepatitis A virus; vaccine; attenuated.

OS Hepatitis A virus, strain HM-175.

XX Key Location/Qualifiers

FT Region 1..23

FT /label=VP4 = 1A

FT Region 24..245

FT /label=VP2 = 1B

FT Region 246..491

FT /label=VP3 = 1C

FT Region 492..791

FT /label=VP1 = 1D

FT Region 792..980

FT /label=2A

FT Region 981..1087

FT /label=2B

FT Region 1088..1422

FT /label=2C

FT Region 1423..1496

FT /label=3A

FT Region 1497..1519

FT /label=3B = VPg

FT Region 1520..1738

FT /label=3C

FT Region 1739..2227

FT /label=3D

XX US4894228-A.

XX 16-JAN-1990.

XX 12-JUL-1988; 88US-0217824.

XX 12-JUL-1988; 88US-0217824.

XX 12-JUL-1988; 88US-0652967.

XX (USSH) US DEPT HEALTH & HUMAN.

XX Purcell RH, Ticehurst JR, Cohen I, Emerson SU, Feinstein SM;

PI Daemer RJ, Gust ID;

XX WPI: 1990-075557/10.

XX N-PSDB; Q03512.

PT Vaccine against hepatitis A virus infection - comprises novel
 XX attenuated hepatitis A virus strain.

PS Claim 1; Fig 1; 18pp; English.

XX The attenuated HAV is useful for inducing protective immunity against
 CC HAV. This strain (pass 35) differs from the wild type HAV HM-175 by
 CC several nucleotide changes distributed throughout the genome, is
 CC attenuated for chimpanzees, elicits serum neutralising antibodies, and is
 CC suitable for use as an HAV vaccine. It is noted that not all the changes
 CC are necessary for attenuation and use as a vaccine.

XX Sequence 2227 AA;

Query Match 100.0%; Score 102; DB 11; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 6.6e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVKGKRLKYAOEE 20
 |||||
 Db 799 pykelrlevkgkrlkyaqee 818

RESULT 9

ID W34074 standard; Protein; 2227 AA.

XX W34074;

XX 27-APR-1998 (first entry)

XX Hepatitis A virus HM-175 protein sequence.

XX HAV; attenuation; 2C protein; 2C chimeric hepatitis A virus;
 XX infection; vaccine.

XX Hepatitis A virus HM-175.

Key Location/Qualifiers

FT Protein 1..23

FT /label= VP4

FT Protein 24..245

FT /label= VP2

FT Protein 246..491

FT /label= VP3

FT Protein 492..791

FT /label= VP1

FT Protein 792..980

FT /label= 2A

FT Protein 981..1087

FT /label= 2B

FT Protein 1088..1422

FT /label= 2C

FT Protein 1423..1496

FT /label= 3A

FT Protein 1497..1519

FT /label= 3B

FT Protein 1520..1738

FT /label= 3C

FT Protein 1739..2227

FT /label= 3D

XX WO9740166-A2.

XX 30-OCT-1997.

XX 18-APR-1997; 97WO-US06506.

XX 19-APR-1996; 96US-0015642.

XX (USSH) US SEC DEPT HEALTH.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA

XX

PI

XX

DR

DR

XX

PT

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

Sequence 2227 AA;

Query Match 100.0%; Score 102; DB 18; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 6.6e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVKGKRLKYAOEE 20

|||||

Db 799 pykelrlevkgkrlkyaqee 818

RESULT 10

BI8607

ID BI8607 standard; Protein; 2227 AA.

XX AC

XX BI8607;

XX DT

XX 15-JAN-2001 (first entry)

XX DE

XX Amino acid sequence of wild type Hepatitis A virus strain HM-175.

XX KW

XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.

XX OS

XX Hepatitis A virus.

XX PN

XX US6113912-A.

XX PD

XX 05-SEP-2000.

XX PF

XX 07-JUN-1995; 95US-0475886.

XX PR

XX 18-SEP-1992; 92US-0947338.

XX PR

XX 17-SEP-1993; 93WO-US08610.

XX PR

XX 10-MAR-1995; 95US-0397232.

XX XX

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PA

XX D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;

XX WPI; 2000-586464/55.

XX N-PSDB; A75476.

XX Novel live hepatitis A virus adapted to growth in human fibroblast cell

PT line useful as vaccine for protecting humans against hepatitis A virus

PT infection, has modified genome compared to wild type -
 XX PS
 XX Disclosure; Fig 6A-K; 72pp; English.

XX The present sequence is derived from a wild type hepatitis A virus
 CC (HAV) strain HM-174. The sequence is modified to produce HAV which
 CC are adapted to growth in the human fibroblast-like cell line MRC-5.
 CC The HAV is able to propagate in MRC-5 cells and retain appropriate
 CC attenuation. It is useful as a live vaccine for prophylaxis of
 CC hepatitis A in humans and other primates.

XX SQ Sequence 2227 AA;

Query Match 100.0%; Score 102; DB 21; Length 2227;

Best Local Similarity 100.0%; Pred. No. 6.6e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PYKELRLEVKGKRLKYAQEE 20
 Db 799 pykelrlevkgkrlkyaqee 818
 |||||

RESULT 11

BI8608
 ID BI8608 standard; Protein; 2227 AA.

XX AC BI8608;

XX DT 15-JAN-2001 (first entry)

DE Amino acid sequence of passage 35 Hepatitis A virus called P-35.

XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
 KW P-35 virus.

XX OS Hepatitis A virus.

XX PN US6113912-A.

XX PD 05-SEP-2000.

XX PF 07-JUN-1995; 95US-0475886.

PR 18-SEP-1992; 92US-0947338.

PR 17-SEP-1993; 93WO-US08610.

PR 10-MAR-1995; 95US-0397232.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
 XX WPI; 2000-586464/55.
 DR N-PSDB; A75477.

XX Novel live hepatitis A virus adapted to growth in human fibroblast cell
 PT line useful as vaccine for protecting humans against hepatitis A virus
 PT infection, has modified genome compared to wild type -

XX Disclosure; Columns 67-78; 72pp; English.

XX The present sequence is derived from passage 35 of a wild type
 CC hepatitis A virus (HAV) strain HM-174. The resulting virus is
 CC designated P-35 virus. The sequence is modified to produce HAV which
 CC are adapted to growth in the human fibroblast-like cell line MRC-5.
 CC The HAV is able to propagate in MRC-5 cells and retain appropriate
 CC attenuation. It is useful as a live vaccine for prophylaxis of
 CC hepatitis A in humans and other primates.

XX SQ Sequence 2227 AA;

Query Match 100.0%; Score 102; DB 21; Length 2227;

Best Local Similarity 100.0%; Pred. No. 6.6e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PYKELRLEVKGKRLKYAQEE 20
 Db 799 pykelrlevkgkrlkyaqee 818
 |||||

RESULT 12

BI8609
 ID BI8609 standard; Protein; 2227 AA.

XX AC BI8609;

XX DT 15-JAN-2001 (first entry)

DE Amino acid sequence of live attenuated Hepatitis A virus 4380.

XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
 KW HAV 4380.

XX OS Hepatitis A virus.

XX PN US6113912-A.

XX PD 05-SEP-2000.

XX PF 07-JUN-1995; 95US-0475886.

PR 18-SEP-1992; 92US-0947338.

PR 17-SEP-1993; 93WO-US08610.

PR 10-MAR-1995; 95US-0397232.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
 XX WPI; 2000-586464/55.
 DR N-PSDB; A75478.

XX Novel live hepatitis A virus adapted to growth in human fibroblast cell
 PT line useful as vaccine for protecting humans against hepatitis A virus
 PT infection, has modified genome compared to wild type -

XX Disclosure; Columns 93-104; 72pp; English.

XX The present sequence is derived from a live attenuated hepatitis A
 CC virus (HAV) of the invention, designated HAV 4380. The sequence is
 CC produced by modifying wild type HAV strain HM-174. The HAV of the
 CC invention are adapted to growth in the human fibroblast-like cell
 CC line MRC-5. The HAV is able to propagate in MRC-5 cells and retain
 CC appropriate attenuation. It is useful as a live vaccine for prophylaxis
 CC of hepatitis A in humans and other primates.

XX SQ Sequence 2227 AA;

Query Match 100.0%; Score 102; DB 21; Length 2227;

Best Local Similarity 100.0%; Pred. No. 6.6e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PYKELRLEVKGKRLKYAQEE 20
 Db 799 pykelrlevkgkrlkyaqee 818
 |||||

RESULT 13

P60066
 ID P60066 standard; Protein; 2227 AA.

XX AC P60066;

XX DT 26-JUN-1991 (first entry)

XX Sequence of viral L434 polypeptide encoded by the complete
DE nucleotide sequence of the HAV genome.

XX Diagnosis; vaccine; passive immunotherapy.

XX Hepatitis A virus.

XX Key Location/Qualifiers

XX Region 1..245

XX /label= p1.1A

XX Region 246..491

XX /label= 1B

XX Region 492..836

XX /label= 1C

XX Region 837..980

XX /label= p2.2A

XX Region 981..1076

XX /label= 2B

XX Region 1077..1422

XX /label= 2C

XX Region 1423..1484

XX /label= p3.3A

XX Region 1485..1507

XX /label= 3B

XX Region 1508..1678

XX /label= 3C

XX Region 1679..2227

XX /label= 3D

XX EP199480-A.

XX 29-OCT-1986.

XX 03-APR-1986; 86EP-0302465.

XX 03-APR-1985; 85US-0719329.

XX (CHIR-) CHIRON CORP.

XX Dina D, Potter SJ, Vannest GA, Caput D;

XX WPI; 1986-286213/44.

XX N-PSDB; N60080.

XX Hepatitis A virus nucleotide sequence and polypeptide - and use

XX in prodn. of vaccines and diagnostic probes

XX Claim 5; Fig 1; 18pp; English.

XX N60080 and oligonucleotide fragments are useful in detection of

XX hepatitis A virus; transformed hosts may be used for expression of

XX polypeptides and fragments useful in vaccines without risk of

XX infection by the virus or in prodn. of particles which are capable

XX of inducing immunocompetent B cells for passive immunotherapy. Pref.

XX epitope is derived from AAs 445-657 or 792-848 of the HAV

XX polypeptide sequence (p60066).

XX Sequence 2227 AA;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX RI5629;

XX 17-MAR-1992 (first entry)

XX Capsid region of cyno-HAV isolate CY-145.

XX Hepatitis A virus; cynomolgus; HAV; monkey; vaccine; macaque.

XX Cynomolgus monkey hepatitis A virus, isolate CY-145.

XX Key Location/Qualifiers

XX Protein 1..245

XX /label= VP0

XX Protein 246..491

XX /label= VP3

XX Protein 492..791

XX /label= VP1

XX Protein 792

XX /label= p2

XX /note= "incomplete"

XX Cleavage-site 245..246

XX Cleavage-site 491..492

XX Cleavage-site 791..792

XX Active-site 315

XX Active-site 593

XX US7678828-A.

XX 12-NOV-1991.

XX 03-APR-1991; 91US-0678828.

XX 03-APR-1991; 91US-0678828.

XX (USSH) US DEPT HEALTH & HUMAN.

XX Hainan OV, Margolis HS, Robertson BH, Brinton MH, Ebert JM;

XX WPI; 1991-376737/51.

XX N-PSDB; Q15180.

XX Hepatitis A virus isolates and DNA - used to prepare vaccines for

XX preventing hepatitis A virus infection.

XX Disclosure; Fig 3; 23pp; English.

XX The sequence was deduced from the nucleotide sequence obtd. by PCR

XX amplification of cyno-HAV viral RNA obtd. from the stool of a

XX cynomolgus monkey with serologically and histologically confirmed

XX spontaneous hepatitis A. The sequence differs from the human HAV

XX isolate HMI75 (Cohen, J.I., et al. (1987) Proc. Natl. Acad. Sci.

XX USA 84, 2497-2501), mainly in the VP3 and VP1 proteins. The Gln-Val

XX pair at the VP3-VP1 cleavage site in the human isolate is replaced

XX by a Gln-Thr pair in the cyno-HAV. The other two cleavage sites are

XX the same. Two residues have been identified as part of the immuno-

XX dominant region (see feature table) and are different to those in

XX the same position in human HAV. The protein and peptides derived

XX from it can be used in the prepn. of vaccines for the prevention of

XX HAV infection.

XX See also RI5056.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Query Match 92.2%; Score 94; DB 12; Length 839;

Best Local Similarity 90.0%; Pred. No. 4.5e-07;

Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20

Db 798 pykelrlvgkqrkyaree 817

RESULT 14

RI5629

ID RI5629 standard; Protein; 839 AA.

```

RESULT 15
W42922
ID W42922 standard; peptide; 20 AA.
XX
AC W42922;
XX
DT 28-APR-1998 (first entry)
XX
DE Immunogenic Hepatitis A virus peptide YK-1315.
XX
KW Immunogenic peptide; immunogenic epitope; P2A protein;
immune response; antibody.
XX
OS Synthetic.
OS Hepatitis A virus.
XX
PN W09740147-A1.
XX
PD 30-OCT-1997.
XX
PF 18-APR-1997; 97WO-US06891.
XX
PR 19-APR-1996; 96US-0015644.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Fields HA, Khudyakov YE;
XX
DR WPI; 1997-535831/49.
XX
PT Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an
immune response to HAV in a mammal or to detect the presence of
antibodies against HAV in a mammal
XX
PS Claim 18; Page 112; 140pp; English.
XX
CC Peptides W42922-30 are immunogenic peptides corresponding to immunogenic
epitopes of the Hepatitis A virus (HAV). The peptides are substantially
similar to a portion of the amino acid sequence of the P2A protein of HAV
corresponding to amino acids 792-980. The present peptide is derived
from amino acids 792-811, and has a reactivity of 54.2% with acute sera.
Compositions containing the peptides can be used to induce an immune
response to HAV in a mammal. The peptides can also be used to detect the
presence of antibodies against HAV in mammalian serum. The peptides can
also be used to make an antibody against HAV by administering the peptide
to a mammal.
XX
SQ Sequence 20 AA;

```

```

Query Match 65.7%; Score 67; DB 18; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 PYKELRLVGVKQR 13
Db 8 pykelrlvvgkqr 20

```

```

Search completed: April 24, 2001, 16:23:14
Job time: 132 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2001, 16:34:58 ; Search time 62.1 Seconds
(without alignments)
6.187 Million cell updates/sec

Title: US-09-171-432A-40

Perfect score: 102

Sequence: 1 PYKELRLEVGKQRLKYAEE 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	102	100.0	1091	6	Patent No. 5516630
2	102	100.0	2227	4	Sequence 2, Appli
3	102	100.0	2227	4	Sequence 4, Appli
4	102	100.0	2227	4	Sequence 6, Appli
5	102	100.0	2227	4	Sequence 2, Appli
6	102	100.0	2227	4	Sequence 4, Appli
7	94	92.2	839	1	Sequence 2, Appli
8	41	40.2	68	1	Sequence 2, Appli
9	41	40.2	68	1	Sequence 4, Appli
10	41	40.2	68	2	Sequence 2, Appli
11	41	40.2	68	2	Sequence 4, Appli
12	41	40.2	69	5	Sequence 5, Appli
13	41	40.2	69	5	Sequence 7, Appli
14	41	40.2	147	1	Sequence 2, Appli
15	41	40.2	147	1	Sequence 3, Appli
16	41	40.2	147	4	Sequence 2, Appli
17	40	39.2	1463	1	Sequence 3, Appli
18	40	39.2	1463	4	Sequence 3, Appli
19	39	38.2	1408	1	Sequence 2, Appli
20	39	38.2	1898	1	Sequence 94, Appl
21	39	38.2	1898	4	Sequence 94, Appl
22	38.5	37.7	641	4	Sequence 160, App
23	38	37.3	233	3	Sequence 8, Appli
24	38	37.3	256	2	Sequence 2, Appli
25	38	37.3	256	4	Sequence 2, Appli
26	37.5	36.8	18	3	Sequence 229, App
27	37.5	36.8	18	3	Sequence 229, App

28	37.5	36.8	18	3	US-08-940-096-229	Sequence 229, App
29	37	36.3	67	3	US-09-120-365-93	Sequence 93, Appl
30	37	36.3	68	1	US-08-606-789-8	Sequence 8, Appli
31	37	36.3	68	2	US-09-111-348-8	Sequence 8, Appli
32	37	36.3	69	5	PCT-US95-06406A-6	Sequence 6, Appli
33	37	36.3	284	2	US-08-766-439-32	Sequence 32, Appl
34	37	36.3	284	2	US-08-766-439-41	Sequence 41, Appl
35	37	36.3	284	2	US-08-766-439-42	Sequence 42, Appl
36	37	36.3	284	2	US-08-766-439-43	Sequence 43, Appl
37	37	36.3	284	2	US-08-766-439-44	Sequence 44, Appl
38	37	36.3	284	2	US-08-766-439-45	Sequence 45, Appl
39	37	36.3	458	1	US-08-336-618-24	Sequence 24, Appl
40	37	36.3	552	3	US-09-120-365-5	Sequence 5, Appli
41	37	36.3	599	3	US-08-556-419-22	Sequence 22, Appl
42	37	36.3	629	3	US-08-556-419-23	Sequence 23, Appl
43	37	36.3	637	2	US-08-426-125-10	Sequence 10, Appl
44	37	36.3	637	2	US-08-455-355-10	Sequence 10, Appl
45	37	36.3	1091	3	US-08-633-768A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
5516630-2
; Patent No. 5516630
; APPLICANT: TICEHURST, JOHN R.; BALTIMORE, DAVID; FEINSTONE,
; STEPHEN M.; PURCELL, ROBERT H.; RACANELLO, VINCENT R.;
; BAROUDY, BAHIGE M.
; TITLE OF INVENTION: METHODS OF DETECTING HEPATITIS A VIRUS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/788,262
; FILING DATE: 06-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 256,135
; FILING DATE: 06-OCT-1988
; APPLICATION NUMBER: 654,942
; FILING DATE: 27-SEP-1984
; APPLICATION NUMBER: 537,911
; FILING DATE: 30-SEP-1983
; SEQ ID NO: 2:
; LENGTH: 1091
5516630-2

Query Match 100.0%; Score 102; DB 6; Length 1091;
Best Local Similarity 100.0%; Pred. No. 8.9e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAEE 20
|||||
DB 1036 PYKELRLEVGKQRLKYAEE 1055

RESULT 2
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2

Query Match      100.0%; Score 102; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PYKELRLEVKGKRLKYAQEE 20
Db 799 PYKELRLEVKGKRLKYAQEE 818
|||||

RESULT 3
US-08-475-886-4
; Sequence 4, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1993-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4

Query Match      100.0%; Score 102; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PYKELRLEVKGKRLKYAQEE 20
Db 799 PYKELRLEVKGKRLKYAQEE 818
|||||

RESULT 4
US-08-475-886-6
; Sequence 6, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1993-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6

Query Match      100.0%; Score 102; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PYKELRLEVKGKRLKYAQEE 20
Db 799 PYKELRLEVKGKRLKYAQEE 818
|||||

RESULT 5
US-08-397-232-2
; Sequence 2, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2

Query Match      100.0%; Score 102; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PYKELRLEVKGKRLKYAQEE 20
Db 799 PYKELRLEVKGKRLKYAQEE 818
|||||

RESULT 6
US-08-397-232-4
; Sequence 4, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4
```

Query Match 100.0%; Score 102; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
|||||
Db 799 PYKELRLEVGKQRLKYAQEE 818

RESULT 7
US-08-087-016-2
; Sequence 2, Application US/08087016
; Patent No. 5430135
; GENERAL INFORMATION:
; APPLICANT: NAINAN, OMANA V.
; APPLICANT: MARGOLIS, HAROLD S.
; APPLICANT: ROBERTSON, BETTY H.
; APPLICANT: BRINTON, MARGO H.
; APPLICANT: EBERT, JAMES W.
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1615 L Street N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,016
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/678,828
; FILING DATE: 03-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: SCOTT, WATSON T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: WTS/5683/83834/SRL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 839 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-087-016-2

Query Match 92.2%; Score 94; DB 1; Length 839;
Best Local Similarity 90.0%; Pred. No. 1.4e-07;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
|||||
Db 798 PYKELRLEVGKQRLKYAQEE 817

RESULT 8
US-08-606-789-2
; Sequence 2, Application US/08606789
; Patent No. 5783418
; GENERAL INFORMATION:

; APPLICANT: Au-Young, Janice
; APPLICANT: Stuart, Susan G.
; APPLICANT: Murry, Lynn E.
; APPLICANT: Guegler, Kark J.
; APPLICANT: Seilhamer, Jeffrey J.
; TITLE OF INVENTION: HUMAN HOMOLOG OF THE RAT G PROTEIN
; TITLE OF INVENTION: GAMMA-5 SUBUNIT
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/606,789
; FILING DATE: Filed Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/440,743
; FILING DATE: May 5, 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/320,011
; FILING DATE: October 5, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Luther, Barbara J.
; REGISTRATION NUMBER: 33,954
; REFERENCE/DOCKET NUMBER: PF-0055 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-852-0195
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 68 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: PITUITARY
; CLONE: 112530
US-08-606-789-2

Query Match 40.2%; Score 41; DB 1; Length 68;
Best Local Similarity 50.0%; Pred. No. 4.6;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 KELRLRLEVGKQRLKYAQ 18
:|||||
Db 15 QQLRLRLEAGLNRVKVSQ 30

RESULT 9
US-08-606-789-4
; Sequence 4, Application US/08606789
; Patent No. 5783418
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Stuart, Susan G.
; APPLICANT: Murry, Lynn E.
; APPLICANT: Guegler, Kark J.
; APPLICANT: Seilhamer, Jeffrey J.
; TITLE OF INVENTION: HUMAN HOMOLOG OF THE RAT G PROTEIN
; TITLE OF INVENTION: GAMMA-5 SUBUNIT
; NUMBER OF SEQUENCES: 31

```

CORRESPONDENCE ADDRESS:
  ADDRESSEE: Incyte Pharmaceuticals, Inc.
  STREET: 3174 Porter Drive
  CITY: Palo Alto
  STATE: CA
  COUNTRY: U.S.
  ZIP: 94303
COMPUTER READABLE FORM:
  MEDIUM TYPE: Diskette
  COMPUTER: IBM Compatible
  OPERATING SYSTEM: DOS
  SOFTWARE: FastSeq Version 1.5
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/606,789
  FILING DATE: Filed Herewith
  CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08/440,743
  FILING DATE: May 5, 1995
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08/320,011
  FILING DATE: October 5, 1994
  ATTORNEY/AGENT INFORMATION:
  NAME: Luther, Barbara J.
  REGISTRATION NUMBER: 33,954
  REFERENCE/DOCKET NUMBER: PF-0055 US
  TELEPHONE: 415-855-0555
  TELEFAX: 415-852-0195
  TELEX:
  INFORMATION FOR SEQ ID NO: 4:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 68 amino acids
  TYPE: amino acid
  STRANDEDNESS: single
  TOPOLOGY: linear
  MOLECULE TYPE: peptide
  IMMEDIATE SOURCE:
  LIBRARY: GenBank
  CLONE: 204241
US-08-606-789-4

```

```

Query Match      40.2%; Score 41; DB 1; Length 68;
Best Local Similarity 50.0%; Pred. No. 4.6;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```

```

Qy      3  KELRLEVGKQRLKYAQ 18
Db      15  QQLRLEAGLNVRKVSQ 30

```

```

RESULT 10
US-09-111-348-2
  Sequence 2, Application US/09111348
  Patent No. 5912130
  GENERAL INFORMATION:
  APPLICANT: Au-Young, Janice
  APPLICANT: Stuart, Susan G.
  APPLICANT: Murry, Lynn E.
  APPLICANT: Guegler, Kark J.
  APPLICANT: Seilhamer, Jeffrey J.
  TITLE OF INVENTION: HUMAN HOMOLOG OF THE RAT G PROTEIN
  NUMBER OF SEQUENCES: 31
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Incyte Pharmaceuticals, Inc.
  STREET: 3174 Porter Drive
  CITY: Palo Alto
  STATE: CA
  COUNTRY: U.S.
  ZIP: 94303
COMPUTER READABLE FORM:

```

```

  MEDIUM TYPE: Diskette
  COMPUTER: IBM Compatible
  OPERATING SYSTEM: DOS
  SOFTWARE: FastSeq Version 1.5
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/111,348
  FILING DATE:
  CLASSIFICATION:
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08/606,789
  FILING DATE:
  APPLICATION NUMBER: 08/440,743
  FILING DATE: May 5, 1995
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08/320,011
  FILING DATE: October 5, 1994
  ATTORNEY/AGENT INFORMATION:
  NAME: Luther, Barbara J.
  REGISTRATION NUMBER: 33,954
  REFERENCE/DOCKET NUMBER: PF-0055 US
  TELECOMMUNICATION INFORMATION:
  TELEPHONE: 415-855-0555
  TELEFAX: 415-852-0195
  TELEX:
  INFORMATION FOR SEQ ID NO: 2:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 68 amino acids
  TYPE: amino acid
  STRANDEDNESS: single
  TOPOLOGY: linear
  MOLECULE TYPE: peptide
  IMMEDIATE SOURCE:
  LIBRARY: PITUITARY
  CLONE: 112530
US-09-111-348-2

```

```

Query Match      40.2%; Score 41; DB 2; Length 68;
Best Local Similarity 50.0%; Pred. No. 4.6;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```

```

Qy      3  KELRLEVGKQRLKYAQ 18
Db      15  QQLRLEAGLNVRKVSQ 30

```

```

RESULT 11
US-09-111-348-4
  Sequence 4, Application US/09111348
  Patent No. 5912130
  GENERAL INFORMATION:
  APPLICANT: Au-Young, Janice
  APPLICANT: Stuart, Susan G.
  APPLICANT: Murry, Lynn E.
  APPLICANT: Guegler, Kark J.
  APPLICANT: Seilhamer, Jeffrey J.
  TITLE OF INVENTION: HUMAN HOMOLOG OF THE RAT G PROTEIN
  NUMBER OF SEQUENCES: 31
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Incyte Pharmaceuticals, Inc.
  STREET: 3174 Porter Drive
  CITY: Palo Alto
  STATE: CA
  COUNTRY: U.S.
  ZIP: 94303
COMPUTER READABLE FORM:
  MEDIUM TYPE: Diskette
  COMPUTER: IBM Compatible
  OPERATING SYSTEM: DOS
  SOFTWARE: FastSeq Version 1.5
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/111,348

```

```

; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/606,789
; FILING DATE:
; APPLICATION NUMBER: 08/440,743
; FILING DATE: May 5, 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/320,011
; FILING DATE: October 5, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Luther, Barbara J.
; REGISTRATION NUMBER: 33,954
; REFERENCE/DOCKET NUMBER: PF-0055 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-852-0195
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 68 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 204241
US-09-111-348-4

```

Query Match 40.2%; Score 41; DB 2; Length 68;
 Best Local Similarity 50.0%; Pred. No. 4.6;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```

QY 3 KELRLEVGKQRLKYAQ 18
Db 15 QQLRLEAGLNRVKVSQ 30

```

```

RESULT 12
PCT-US95-06406A-5
; Sequence 5, Application PC/TUS9506406A
; GENERAL INFORMATION:
; APPLICANT: Janet D. Robishaw, Charles Kunsch
; TITLE OF INVENTION: cDNA Clones Encoding Human G Protein
; TITLE OF INVENTION: Subunits
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE:
; STREET:
; CITY:
; STATE:
; COUNTRY:
; ZIP:
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
; COMPUTER: IBM 486
; OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/06406A
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:

```

```

; TELEFAX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 69
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
PCT-US95-06406A-5

```

Query Match 40.2%; Score 41; DB 5; Length 69;
 Best Local Similarity 50.0%; Pred. No. 4.7;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```

QY 3 KELRLEVGKQRLKYAQ 18
Db 15 QQLRLEAGLNRVKVSQ 30

```

```

RESULT 13
PCT-US95-06406A-7
; Sequence 7, Application PC/TUS9506406A
; GENERAL INFORMATION:
; APPLICANT: Janet D. Robishaw, Charles Kunsch
; TITLE OF INVENTION: cDNA Clones Encoding Human G Protein
; TITLE OF INVENTION: Subunits
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE:
; STREET:
; CITY:
; STATE:
; COUNTRY:
; ZIP:
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
; COMPUTER: IBM 486
; OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/06406A
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:
; TELEFAX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 69
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
PCT-US95-06406A-7

```

Query Match 40.2%; Score 41; DB 5; Length 69;
 Best Local Similarity 43.8%; Pred. No. 4.7;
 Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

```

QY 3 KELRLEVGKQRLKYAQ 18
Db 15 EQLKLEAGVERIKVSQ 30

```

```

RESULT 14
US-08-264-003B-2
; Sequence 2, Application US/08264003B
; Patent No. 5696076

```

GENERAL INFORMATION:
APPLICANT: Gentz, Reiner
TITLE OF INVENTION: 5-LIPOXYGENASE-ACTIVATING PROTEIN II
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELIA, BYRNE, BAIN, GILFILLAN, CECCHI,
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NJ
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/264,003B
FILING DATE: 06-JUN-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ferraro, Gregory D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-117
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 147 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-264-003B-2

Query Match 40.2%; Score 41; DB 1; Length 147;
Best Local Similarity 80.0%; Pred. No. 10;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 7 LEVGKQRLKY 16
I:|||||
Db 24 LQVGKARLKY 33

RESULT 15
US-08-959-865-3
Sequence 3, Application US/08959865
Patent No. 5919627
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: MICROSOMAL GLUTATHIONE-S TRANSFERASE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/959,865
FILING DATE: Herewith
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0411 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 147 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: COLNNOT07
CLONE: 903729
US-08-959-865-3

Query Match 40.2%; Score 41; DB 2; Length 147;
Best Local Similarity 80.0%; Pred. No. 10;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 7 LEVGKQRLKY 16
I:|||||
Db 24 LQVGKARLKY 33

Search completed: April 24, 2001, 16:34:59
Job time: 530 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2001, 16:25:42 ; Search time 73.14 Seconds
(without alignments)
18.792 Million cell updates/sec

Title: US-09-171-432A-40

Perfect score: 102

Sequence: 1 PYKELRLEVGKORLKYAQEE 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR:67.*

1: p1r1.*

2: p1r2.*

3: p1r3.*

4: p1r4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	102	100.0	55	2 P00433	genome polyprotein
2	102	100.0	56	2 P00434	genome polyprotein
3	102	100.0	56	2 P00428	genome polyprotein
4	102	100.0	56	2 P00430	genome polyprotein
5	102	100.0	56	2 P00430	genome polyprotein
6	102	100.0	341	2 S04137	genome polyprotein
7	102	100.0	852	1 GNNYHA	genome polyprotein
8	102	100.0	1358	2 A03905	genome polyprotein
9	102	100.0	2227	1 GNNYHM	genome polyprotein
10	102	100.0	2227	1 GNNYHR	genome polyprotein
11	102	100.0	2227	1 GNNYK	genome polyprotein
12	102	100.0	2227	1 GNNYHB	genome polyprotein
13	100	98.0	56	2 P00427	genome polyprotein
14	99	97.1	56	2 P00429	genome polyprotein
15	95	93.1	56	2 P00436	genome polyprotein
16	95	93.1	2230	1 GNNYSA	genome polyprotein
17	94	92.2	55	2 P00435	genome polyprotein
18	94	92.2	56	2 P00437	genome polyprotein
19	94	92.2	839	1 GNNYS2	genome polyprotein
20	93	91.2	56	2 P00431	genome polyprotein
21	62	60.8	319	2 JH0135	genome polyprotein
22	46	45.1	219	2 T45997	hypothetical prote
23	46	45.1	285	2 T29804	hypothetical prote
24	45	44.1	209	2 D72121	ATP synthase chain
25	44	43.1	290	2 C71701	branched-chain ami
26	44	43.1	776	2 T09938	hypothetical prote
27	43	42.2	208	1 F70485	uracil phosphoribo
28	43	42.2	465	2 T25441	hypothetical prote
29	43	42.2	509	2 G81929	probable iron-upta

30	43	42.2	534	2	JC5096	transposase - fung
31	43	42.2	3587	2	T31075	tyrocidine synthet
32	42.5	41.7	179	2	A65027	hypothetical prote
33	42	41.2	452	2	B83623	probable porin PA0
34	42	41.2	736	2	B82944	ribose/galactose A
35	42	41.2	752	2	A65089	DNA topoisomerase
36	42	41.2	752	2	A45582	DNA topoisomerase
37	42	41.2	939	2	H71532	valine--tRNA ligas
38	42	41.2	4131	2	T21085	hypothetical prote
39	41	40.2	68	2	B42243	GTP-binding regula
40	41	40.2	68	2	I39158	GTP-binding regula
41	41	40.2	151	2	A72409	conserved hypother
42	41	40.2	412	2	E70108	aminopeptidase II
43	41	40.2	446	2	G70774	probable atp prot
44	41	40.2	515	2	H75579	hypothetical prote
45	41	40.2	656	2	E75468	hypothetical prote

ALIGNMENTS

RESULT 1

PQ0433

genome polyprotein - human hepatitis A virus (strain PA21) (fragment)

C:Species: human hepatitis A virus

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C:Accession: PQ0433

R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; W1

J. Gen. Virol. 73, 1365-1377, 1992

A:Title: Genetic relatedness of hepatitis A virus strains recovered from different ge

A:Reference number: PQ0427; MUID:92300330

A:Accession: PQ0433

A:Molecule type: mRNA

A:Residues: 1-55 <ROB>

A>Note: this protein is from the VP1/2A Junction region

C:Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; polyprotein

Query Match 100.0%; Score 102; DB 2; Length 55;

Best Local Similarity 100.0%; Pred. No. 1.3e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PYKELRLEVGKORLKYAQEE 20

|||||

Db 35 PYKELRLEVGKORLKYAQEE 54

RESULT 2

PQ0434

genome polyprotein - human hepatitis A virus (strain KPH) (fragment)

C:Species: human hepatitis A virus

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C:Accession: PQ0434

R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; W1

J. Gen. Virol. 73, 1365-1377, 1992

A:Title: Genetic relatedness of hepatitis A virus strains recovered from different ge

A:Reference number: PQ0427; MUID:92300330

A:Accession: PQ0434

A:Molecule type: mRNA

A:Residues: 1-56 <ROB>

A>Note: this protein is from the VP1/2A Junction region

C:Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; polyprotein

Query Match 100.0%; Score 102; DB 2; Length 56;

Best Local Similarity 100.0%; Pred. No. 1.3e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PYKELRLEVGKORLKYAQEE 20

|||||

Db 36 PYKELRLEVGKORLKYAQEE 55

```

RESULT 3
PQ0428
genome polyprotein - human hepatitis A virus (strain TKM002) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0428
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A:Reference number: PQ0427; MUID:92300330
A:Accession: PQ0428
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A:Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 100.0%; Score 102; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PYKELRLEVGKQRLKYAQEE 20
Db 36 PYKELRLEVGKQRLKYAQEE 55
|||||

RESULT 4
PQ0432
genome polyprotein - human hepatitis A virus (strain CF-53) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0432
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A:Reference number: PQ0427; MUID:92300330
A:Accession: PQ0432
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A:Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 100.0%; Score 102; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PYKELRLEVGKQRLKYAQEE 20
Db 36 PYKELRLEVGKQRLKYAQEE 55
|||||

RESULT 5
PQ0430
genome polyprotein - human hepatitis A virus (strain S23-1) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0430
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A:Reference number: PQ0427; MUID:92300330
A:Accession: PQ0430
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A:Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 100.0%; Score 102; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PYKELRLEVGKQRLKYAQEE 20
Db 36 PYKELRLEVGKQRLKYAQEE 55
|||||
```

```

Query Match 100.0%; Score 102; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PYKELRLEVGKQRLKYAQEE 20
Db 36 PYKELRLEVGKQRLKYAQEE 55
|||||

RESULT 6
S04137
genome polyprotein - human hepatitis A virus (strain LCDC-1) (fragment)
C:Species: human hepatitis A virus
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 21-Jul-2000
C:Accession: S04137
R:Andonov, A.P.; Lau, P.C.K.; Chaudhary, R.
Nucleic Acids Res. 17, 3594, 1989
A:Title: Nucleotide sequence of the VP1 gene from a Chinese strain of hepatitis A vir
A:Reference number: S04137; MUID:89263805
A:Accession: S04137
A:Molecule type: mRNA
A:Residues: 1-341 <AND>
A:Cross-references: EMBL:X14666; NID:g62301; PIDN:CAA32794.1; PID:g4377576
C:Genetics:
C:Gene: VP1
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; polyprotein
F:2-340/Product: coat protein ID (VP1) #status predicted <MAT>

Query Match 100.0%; Score 102; DB 2; Length 341;
Best Local Similarity 100.0%; Pred. No. 9.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PYKELRLEVGKQRLKYAQEE 20
Db 303 PYKELRLEVGKQRLKYAQEE 322
|||||

RESULT 7
GNNYHA
genome polyprotein - human hepatitis A virus (strain CR326) (fragment)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C:Accession: A03904
R:Linemeyer, D.L.; Menke, J.G.; Martin-Gallardo, A.; Hughes, J.V.; Young, A.; Mitra,
J. Virol. 54, 247-255, 1985
A:Title: Molecular cloning and partial sequencing of hepatitis A viral cDNA.
A:Reference number: A03904; MUID:85185648
A:Accession: A03904
A:Molecule type: genomic RNA
A:Residues: 1-852 <LIN>
A:Cross-references: EMBL:M10033; NID:g329592; PIDN:AAA45470.1; PID:g329593
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein
F:1-245/Product: coat protein 1A #status predicted <CIA>
F:246-491/Product: coat protein 1B #status predicted <CLIC>
F:492-836/Product: coat protein 1C #status predicted <CLIC>
F:837-852/Product: core protein 2A (fragment) #status predicted <C2A>

Query Match 100.0%; Score 102; DB 1; Length 852;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PYKELRLEVGKQRLKYAQEE 20
Db 799 PYKELRLEVGKQRLKYAQEE 818
|||||
```

```

RESULT 8
A03905
genome polyprotein (version 2) - human hepatitis A virus (fragments)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 2A; protein
C:Species: human hepatitis A virus
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 15-Nov-1996
C:Accession: A03905
R:Baroudy, B.M.; Ticehurst, J.R.; Miele, T.A.; Maizel Jr., J.V.; Purcell, R.H.; Feinstone
Proc. Natl. Acad. Sci. U.S.A. 82, 2143-2147, 1985
A:Title: Sequence analysis of hepatitis A virus cDNA coding for capsid proteins and RNA
A:Reference number: A03905; MUID:85166289
A:Accession: A03905
A:Molecule type: genomic RNA
A:Residues: 1-1358 <BAK>
C:Superfamily: hepatitis A virus genome polyprotein
F:1-245/Product: coat protein; core protein; polyprotein
F:246-491/Product: coat protein 1A #status predicted <C1A>
F:492-836/Product: coat protein 1B #status predicted <C1B>
F:837-854/Product: coat protein 1C #status predicted <C1C>
F:855-1358/Product: protein 3D (RNA polymerase) (fragment) #status predicted <C3D>

Query Match 100.0%; Score 102; DB 2; Length 1358;
Best Local Similarity 100.0%; Pred. No. 4.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
|||||
DB 799 PYKELRLEVGKQRLKYAQEE 818

RESULT 9
GNNYHM
genome polyprotein - human hepatitis A virus (strain HM-175, wild type)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pro
B: RNA-directed RNA polymerase (BC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999
C:Accession: A25981
R:Cohen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.
J. Virol. 61, 50-59, 1987
A:Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with di
A:Reference number: A25981; MUID:87061253
A:Accession: A25981
A:Molecule type: genomic RNA
A:Residues: 1-2227 <COH>
A:Cross-references: EMBL:M14707; NID:g329582; PIDN:AAA45465.1; PID:g329583
C:Superfamily: hepatitis A virus genome polyprotein
F:1-23/Product: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltrans
F:24-245/Product: coat protein 1A #status predicted <VP4>
F:246-491/Product: coat protein 1B #status predicted <VP2>
F:492-791/Product: coat protein 1C #status predicted <VP3>
F:792-980/Product: coat protein 1D #status predicted <VP1>
F:981-1087/Product: core protein 2A #status predicted <C2A>
F:1088-1422/Product: core protein 2B #status predicted <C2B>
F:1423-1496/Product: core protein 2C #status predicted <C2C>
F:1497-1519/Product: protein 3A #status predicted <C3A>
F:1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3B>
F:1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 102; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 7.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
|||||
DB 799 PYKELRLEVGKQRLKYAQEE 818

```

```

RESULT 10
GNMYHR
genome polyprotein - human hepatitis A virus
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 2A; core
NA polymerase (BC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C:Accession: A03903
R:Najarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van N
Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985
A:Title: Primary structure and gene organization of human hepatitis A virus.
A:Reference number: A03903; MUID:85190549
A:Accession: A03903
A:Molecule type: genomic RNA
A:Residues: 1-2227 <NAJ>
A:Cross-references: GB:K02990; NID:g329596; PIDN:AAA45472.1; PID:g329597
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltr
F:1-245/Product: coat protein 1A #status predicted <C1A>
F:246-491/Product: coat protein 1B #status predicted <C1B>
F:492-836/Product: coat protein 1C #status predicted <C1C>
F:837-980/Product: core protein 2A #status predicted <C2A>
F:981-1076/Product: core protein 2B #status predicted <C2B>
F:1077-1422/Product: core protein 2C #status predicted <C2C>
F:1423-1484/Product: protein 3A #status predicted <C3A>
F:1485-1507/Product: protein 3B #status predicted <C3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 102; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 7.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
|||||
DB 799 PYKELRLEVGKQRLKYAQEE 818

RESULT 11
GNMYMK
genome polyprotein - human hepatitis A virus (strain HM-175/7MK-5, attenuated HAV)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core
NA polymerase (BC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999
C:Accession: A94149; A25914; A94508
R:Cohen, J.I.; Rosenblum, B.; Ticehurst, J.R.; Daemer, R.J.; Feinstein, S.M.; Purcell
Proc. Natl. Acad. Sci. U.S.A. 84, 2497-2501, 1987
A:Title: Complete nucleotide sequence of an attenuated hepatitis A virus: comparison
A:Reference number: A94149; MUID:87175701
A:Accession: A94149
A:Status: nucleic acid sequence not shown
A:Molecule type: genomic RNA
A:Residues: 1-2227 <COH>
A:Cross-references: EMBL:M16632; NID:g329594; PIDN:AAA45471.1; PID:g329595
A:Note: submitted to GenBank, August 1987
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltr
F:1-245/Product: coat protein 1A #status predicted <P1A>
F:246-491/Product: coat protein 1B #status predicted <P1B>
F:492-836/Product: coat protein 1C #status predicted <P1C>
F:837-980/Product: core protein 2A #status predicted <P2A>
F:981-1076/Product: core protein 2B #status predicted <P2B>
F:1077-1422/Product: core protein 2C #status predicted <P2C>
F:1423-1484/Product: protein 3A #status predicted <P3A>
F:1485-1507/Product: protein 3B #status predicted <P3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

```

```
Query Match      100.0%; Score 102; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 7.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVKGKORLKYAQEE 20
Db 799 PYKELRLEVKGKORLKYAQEE 818

RESULT 12
GNVYHB
genome polyprotein - human hepatitis A virus (strain MBB)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein 3D
Vpg: protein 3A; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-May-1996
R:Paul, A.V.; Tada, H.; von der Helm, K.; Wissel, T.; Kiehn, R.; Wimmer, E.; Deinhardt,
Virus Res. 8, 153-171, 1987
A:Title: The entire nucleotide sequence of the genome of human hepatitis A virus (isolat
A:Reference number: JS0303; MUID:88045071
A:Accession: JS0303
A:Molecule type: genomic RNA
A:Residues: 1-2227 <PAU>
A:Cross-references: EMBL:M20273
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; hydr
F:1-23/Product: coat protein 1A #status predicted <VP4>
F:24-246/Product: coat protein 1B #status predicted <VP3>
F:247-491/Product: coat protein 1C #status predicted <VP3>
F:492-836/Product: coat protein 1D #status predicted <VP1>
F:837-980/Product: core protein 2A #status predicted <P2A>
F:981-1108/Product: core protein 2B #status predicted <P2B>
F:1109-1438/Product: core protein 2C #status predicted <P2C>
F:1439-1496/Product: protein 3A #status predicted <P3A>
F:1497-1519/Product: genome-linked protein vpg #status predicted <VPG>
F:1520-1736/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match      100.0%; Score 102; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 7.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVKGKORLKYAQEE 20
Db 799 PYKELRLEVKGKORLKYAQEE 818

RESULT 13
genome polyprotein - human hepatitis A virus (strain EP-35.730) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0427
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A:Reference number: PQ0427; MUID:92300330
A:Accession: PQ0427
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
C:Comment: This protein is from the VP1/2A Junction region.
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match      98.0%; Score 100; DB 2; Length 56;
Best Local Similarity 95.0%; Pred. No. 2.8e-09;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVKGKORLKYAQEE 20
```

```
Db 36 PYKELRLEVKGKORLKYAQEE 55

RESULT 14
PQ0429
genome polyprotein - human hepatitis A virus (strain PRC16) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0429
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Wi
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different ge
A:Reference number: PQ0427; MUID:92300330
A:Accession: PQ0429
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A:Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match      97.1%; Score 99; DB 2; Length 56;
Best Local Similarity 95.0%; Pred. No. 4.1e-09;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVKGKORLKYAQEE 20
Db 36 PYKELRLEVKGKORLKYAQEE 55

RESULT 15
PQ0436
genome polyprotein - human hepatitis A virus (strain AGM27) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0436
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Wi
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different ge
A:Reference number: PQ0427; MUID:92300330
A:Accession: PQ0436
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A:Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match      93.1%; Score 95; DB 2; Length 56;
Best Local Similarity 90.0%; Pred. No. 1.8e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PYKELRLEVKGKORLKYAQEE 20
Db 36 PYKELRLEVKGKORLKYAMEE 55

Search completed: April 24, 2001, 16:25:46
Job time: 280 sec
```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2001, 16:33:53 ; Search time 38.74 Seconds
(without alignments)
17.685 Million cell updates/sec

Title: US-09-171-432A-40

Perfect score: 102

Sequence: 1 PYKELRLEVGKORLYAQEE 20

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	102	100.0	341	1 POLG_HPAV1	P13672 hepatitis a
2	102	100.0	808	1 POLG_HPAVG	Q02381 hepatitis a
3	102	100.0	852	1 POLG_HPAVC	P08442 hepatitis a
4	102	100.0	2226	1 POLG_HPAV2	P26580 hepatitis a
5	102	100.0	2226	1 POLG_HPAV4	P26581 hepatitis a
6	102	100.0	2226	1 POLG_HPAV8	P26582 hepatitis a
7	102	100.0	2227	1 POLG_HPAVH	P08617 hepatitis a
8	102	100.0	2227	1 POLG_HPAVL	P08641 hepatitis a
9	102	100.0	2227	1 POLG_HPAVM	P13901 hepatitis a
10	95	93.1	2230	1 POLG_HPAVS	P14553 simian hepa
11	94	92.2	839	1 POLG_HPAVT	P31788 simian hepa
12	45	44.1	209	1 VATD_CHLPN	Q02991 chlamydia p
13	44	43.1	290	1 ILVE_RICPR	O05970 rickettsia
14	44	43.1	583	1 FOJO_DROME	P54360 drosophila
15	43	42.2	208	1 UPP_AQUAE	O67914 aquifex aeo
16	43	42.2	507	1 DNAA_RHIME	P33890 rhizobium m
17	43	42.2	3587	1 TYCB_BACBR	O30408 b tyrocidin
18	42	41.2	752	1 PARC_ECOLI	P20082 escherichia
19	42	41.2	752	1 PARC_SALTY	P26973 salmonella
20	42	41.2	939	1 SVV_CHLTR	O84304 chlamydia t
21	41	40.2	68	1 GRG5_HUMAN	P30670 homo sapien
22	41	40.2	68	1 GBGA_HUMAN	P50151 homo sapien
23	41	40.2	147	1 GST2_HUMAN	Q09735 homo sapien
24	41	40.2	318	1 YZ34_AQUAE	O66423 aquifex aeo
25	41	40.2	446	1 ATPD_MYCTU	Q10594 mycobacteri
26	40	39.2	69	1 MOP_HAEIN	P45183 haemophilus
27	40	39.2	456	1 EFIA_DICDI	P18624 dictyosteli
28	40	39.2	470	1 BFRI_YEAST	P38934 saccharomyc
29	40	39.2	723	1 SYM_PYRHO	O58721 pyrococcus
30	40	39.2	733	1 ERG7_RAT	P48450 rattus norv
31	40	39.2	764	1 UBFL_HUMAN	P17480 homo sapien
32	40	39.2	764	1 UBFL_RAT	P25977 rattus norv
33	40	39.2	765	1 UBFL_MOUSE	P25976 mus musculus

34 40 39.2 795 1 LON_MYCPN P78025 mycoplasma
35 40 39.2 939 1 SYV_CHLMU Q9PK91 chlamydia m
36 40 39.2 3859 1 RPOA_LELV Q04561 lelystad vi
37 39.5 38.7 971 1 SN2L_CAEEL P41877 caenorhabdi
38 39.5 38.7 976 1 SN21_HUMAN P28370 homo sapien
39 39.5 38.7 1812 1 BRCL_MOUSE P48754 mus musculus
40 39 38.2 46 1 DIUH_LOCOMI P23465 locustula mig
41 39 38.2 362 1 ILVE_STRCO O86505 streptomyce
42 39 38.2 424 1 EXON_NPVOP P24081 orgyia pseu
43 39 38.2 508 1 V56K_PLRV1 P17525 potato leaf
44 39 38.2 508 1 V56K_PLRVW P11626 potato leaf
45 39 38.2 653 1 PABP_SCHPO P31209 schizosacch

ALIGNMENTS

RESULT 1
POLG_HPAV1
ID POLG_HPAV1 STANDARD; PRT; 341 AA.
AC P13672;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: COAT PROTEINS VP1 TO VP3; CORE PROTEIN P2A] (FRAGMENT).
DE P2A] (FRAGMENT).
OS Hepatitis A virus (strain LCDC-1).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12093;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89263805; Pubmed=2542903;
RA Andonov A.P., Lau P., Chaudhary R.;
RT "Nucleotide sequence of the VP1 gene from a Chinese strain of
hepatitis A virus (HAV)."
RL Nucleic Acids Res. 17:3594-3594(1989).
CC -! SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
VP3, AND VP4.
CC -! PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X14666; CAA32794.1;
DR PIR; S04137; S04137.
KW Polyprotein; Coat protein; Core protein.
FT NON_TER 1 1
FT CHAIN <1 1 COAT PROTEIN VP3 (IC).
FT CHAIN 2 340 COAT PROTEIN VP1 (ID).
FT CHAIN 341 >341 CORE PROTEIN P2A.
FT NON_TER 341 341
SQ SEQUENCE 341 AA; 38003 MW; 066918289BF126D5 CRC64;

Query Match 100.0%; Score 102; DB 1; Length 341;
Best Local Similarity 100.0%; Pred. No. 6.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKORLYAQEE 20

DB 303 PYKELRLEVGKORLYAQEE 322

RESULT 2

POLG_HPAVG STANDARD; PRT; 808 AA.
ID POLG_HPAVG

```
AC Q02381;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: COAT PROTEINS VP1 TO VP4; CORE PROTEIN
DE P2A] (FRAGMENT).
OS Hepatitis A virus (strain GA76).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=31706;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92260183; PubMed=1316423;
RA Khanna B., Spelbring J.E., Innis B.L., Robertson B.H.;
RT "Characterization of a genetic variant of human hepatitis A virus.";
RL J. Med. Virol. 36:118-124(1992).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M66695; AAA45477.1; -.
CC Polyprotein; Coat protein; Core protein.
CC CHAIN 1 2 COAT PROTEIN VP4 (PIA).
CC CHAIN 3 223 COAT PROTEIN VP2 (PIB).
CC CHAIN 224 470 COAT PROTEIN VP3 (PIC).
CC CHAIN 471 770 COAT PROTEIN VP1 (PID).
CC CHAIN 771 >808 CORE PROTEIN P2A.
CC NON_TER 808 808
CC SEQUENCE 808 AA; 90632 MW; D80CE7E57A479C12 CRC64;
CC -----
Query Match 100.0%; Score 102; DB 1; Length 808;
Best Local Similarity 100.0%; Pred. No. 1.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PYKELRLEVGKQRLKYAQEE 20
DB 778 PYKELRLEVGKQRLKYAQEE 797
RESULT 3
POLG_HPAVC
ID POLG_HPAVC STANDARD; PRT; 852 AA.
AC P06442; Q83741; Q83742;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: COAT PROTEINS VP1 TO VP4; CORE PROTEIN
DE P2A] (FRAGMENT).
OS Hepatitis A virus (strain CR326).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12097;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85185648; PubMed=2385793;
RA Linemeyer D.L., Menke J.G., Martin-Gallardo A., Hughes J.V.,
RA Young A., Mitra S.W.;
RT "Molecular cloning and partial sequencing of hepatitis A viral cDNA.";
RL J. Virol. 54:247-255(1985).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
```

```
CC -!- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M10033; AAA45470.1; -.
CC Polyprotein; Coat protein; Core protein.
CC CHAIN 1 23 COAT PROTEIN VP4 (PIA).
CC CHAIN 24 245 COAT PROTEIN VP2 (PIB).
CC CHAIN 246 491 COAT PROTEIN VP3 (PIC).
CC CHAIN 492 836 COAT PROTEIN VP1 (PID).
CC CHAIN 837 >852 CORE PROTEIN P2A.
CC NON_TER 852 852
CC SEQUENCE 852 AA; 95563 MW; 73D3ED0AD532820E CRC64;
CC -----
Query Match 100.0%; Score 102; DB 1; Length 852;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PYKELRLEVGKQRLKYAQEE 20
DB 799 PYKELRLEVGKQRLKYAQEE 818
RESULT 4
POLG_HPAV2
ID POLG_HPAV2 STANDARD; PRT; 2226 AA.
AC P26580;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: COAT PROTEINS VP1 TO VP4; CORE PROTEINS
DE P2A TO P2C; PROBABLE PROTEINS P3A TO P3C; RNA-DIRECTED RNA POLYMERASE
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 24a).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12094;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromeans T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination.";
RL J. Virol. 65:2056-2065(1991).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M59810; AAA45468.1; -.
CC MEROPS; C03.005; -.
CC InterPro; IPR000605; -.
CC InterPro; IPR001205; -.
CC -----
```


DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam: PF00910; RNA_helicase; 1.
 KW Polyprotein; Coat protein; Core protein; Transferase;
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
 FT CHAIN 1 23
 FT CHAIN 24 245
 FT CHAIN 246 491
 FT CHAIN 492 794
 FT CHAIN 795 900
 FT CHAIN 901 1087
 FT CHAIN 1088 1422
 FT CHAIN 1423 1495
 FT CHAIN 1496 1518
 FT CHAIN 1519 1737
 FT CHAIN 1738 2226
 SQ SEQUENCE 2226 AA; 251152 MW; 6CD85A91D6B4E2BF CRC64;
 Query Match 100.0%; Score 102; DB 1; Length 2226;
 Best Local Similarity 100.0%; Pred. No. 4.8e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PYKELRLEVGKORLYAQEE 20
 DB 799 PYKELRLEVGKORLYAQEE 818
 RESULT 5
 POLG_HPAV4
 ID POLG_HPAV4 STANDARD; PRT: 2226 AA.
 AC P26581;
 DT 01-AUG-1992 (Rel. 23; Created)
 DT 01-AUG-1992 (Rel. 23; Last sequence update)
 DT 15-DEC-1998 (Rel. 37; Last annotation update)
 DE GENOME POLYPROTEIN [CONTAINS: COAT PROTEINS VP1 TO VP4; CORE PROTEINS
 DE P2A TO P2C; PROBABLE PROTEINS P3A TO P3C; RNA-DIRECTED RNA POLYMERASE
 DE P3D (EC 2.7.7.48)].
 OS Hepatitis A virus (strain 43c).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12095;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91162758; PubMed=1705995;
 RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
 RA Cromeans T., Jansen R.W.;
 RT "Antigenic and genetic variation in cytopathic hepatitis A virus
 RT variants arising during persistent infection: evidence for genetic
 RT recombination.";
 RL J. Virol. 65:2056-2065(1991).
 CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
 CC VP3, AND VP4.
 CC -!- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -!- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M59809; AAA45469.1; -
 CC MEROPS: C03.005; -
 CC InterPro: IPR000605; -
 CC InterPro: IPR001205; -
 CC Pfam: PF00680; RNA_dep_RNA_pol; 1.
 CC Pfam: PF00910; RNA_helicase; 1.
 KW Polyprotein; Coat protein; Core protein; Transferase;
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
 FT CHAIN 1 23
 FT CHAIN 24 245
 FT CHAIN 246 491
 FT CHAIN 492 794
 FT CHAIN 795 900
 FT CHAIN 901 1087
 FT CHAIN 1088 1422
 FT CHAIN 1423 1495
 FT CHAIN 1496 1518
 FT CHAIN 1519 1737
 FT CHAIN 1738 2226
 SQ SEQUENCE 2226 AA; 251152 MW; 6CD85A91D6B4E2BF CRC64;

FT CHAIN 24 245
 FT CHAIN 246 491
 FT CHAIN 492 794
 FT CHAIN 795 900
 FT CHAIN 901 1087
 FT CHAIN 1088 1422
 FT CHAIN 1423 1495
 FT CHAIN 1496 1518
 FT CHAIN 1519 1737
 FT CHAIN 1738 2226
 SQ SEQUENCE 2226 AA; 251107 MW; 403B4CA80B09BF75 CRC64;
 Query Match 100.0%; Score 102; DB 1; Length 2226;
 Best Local Similarity 100.0%; Pred. No. 4.8e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PYKELRLEVGKORLYAQEE 20
 DB 799 PYKELRLEVGKORLYAQEE 818
 RESULT 6
 POLG_HPAV8
 ID POLG_HPAV8 STANDARD; PRT: 2226 AA.
 AC P26582;
 DT 01-AUG-1992 (Rel. 23; Created)
 DT 01-AUG-1992 (Rel. 23; Last sequence update)
 DT 15-DEC-1998 (Rel. 37; Last annotation update)
 DE GENOME POLYPROTEIN [CONTAINS: COAT PROTEINS VP1 TO VP4; CORE PROTEINS
 DE P2A TO P2C; PROBABLE PROTEINS P3A TO P3C; RNA-DIRECTED RNA POLYMERASE
 DE P3D (EC 2.7.7.48)].
 OS Hepatitis A virus (strain 18f).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12096;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91162758; PubMed=1705995;
 RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
 RA Cromeans T., Jansen R.W.;
 RT "Antigenic and genetic variation in cytopathic hepatitis A virus
 RT variants arising during persistent infection: evidence for genetic
 RT recombination.";
 RL J. Virol. 65:2056-2065(1991).
 CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
 CC VP3, AND VP4.
 CC -!- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -!- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M59808; AAA45467.1; -
 CC MEROPS: C03.005; -
 CC InterPro: IPR000605; -
 CC InterPro: IPR001205; -
 CC Pfam: PF00680; RNA_dep_RNA_pol; 1.
 CC Pfam: PF00910; RNA_helicase; 1.
 KW Polyprotein; Coat protein; Core protein; Transferase;
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
 FT CHAIN 1 23
 FT CHAIN 24 245
 FT CHAIN 246 491
 FT CHAIN 492 794
 FT CHAIN 795 900
 FT CHAIN 901 1087
 FT CHAIN 1088 1422
 FT CHAIN 1423 1495
 FT CHAIN 1496 1518
 FT CHAIN 1519 1737
 FT CHAIN 1738 2226
 SQ SEQUENCE 2226 AA; 251107 MW; 403B4CA80B09BF75 CRC64;

```
FT CHAIN 1088 1422 CORE PROTEIN P2C
FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251292 MW; 24964A63396C8D6B CRC64;

Query Match 100.0%; Score 102; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 4.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PYKELRLEVGVKQRLKYAQEE 20
Db 799 PYKELRLEVGVKQRLKYAQEE 818
|||||

RESULT 7
POLG_HPAVH STANDARD; PRT; 2227 AA.
AC P08617; P06443; Q81082;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: COAT PROTEINS VP1 TO VP4; CORE PROTEINS
DE P2A TO P2C; PROBABLE PROTEINS P3A TO P3C; RNA-DIRECTED RNA POLYMERASE
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain HM-175).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12098;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WILD TYPE;
RX MEDLINE=87061253; PubMed=3023706;
RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A.,
RA Baroudy B.M.;
RT "Complete nucleotide sequence of wild-type hepatitis A virus:
RT comparison with different strains of hepatitis A virus and other
RT picornaviruses.";
RL J. Virol. 61:50-59(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATTENUATED;
RX MEDLINE=87175701; PubMed=3031686;
RA Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstein S.M.,
RA Purcell R.H.;
RT "Complete nucleotide sequence of an attenuated hepatitis A virus:
RT comparison with wild-type virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).
RN [3]
RP SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.
RX MEDLINE=85166289; PubMed=2984684;
RA Baroudy B.M., Ticehurst J.R., Miele T.A., Maizel J.V. Jr.,
RA Purcell R.H., Feinstein S.M.;
RT "Sequence analysis of hepatitis A virus cDNA coding for capsid
RT proteins and RNA polymerase.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 Icosaedral UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- MISCELLANEOUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED
CC FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -!- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT
CC SHOWN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

```
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M14114; AAA45475.1; -
DR EMBL; M14707; AAA45465.1; -
DR EMBL; M14707; AAA45466.1; ALT_INIT.
DR EMBL; M16632; AAA45471.1; -
DR PIR; A25981; GNNYHM.
DR PIR; A25914; GNNYMK.
DR PIR; A03905; A03905.
DR MEROPS; C03.005; -
DR InterPro; IPR001205; -
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1496
FT CHAIN 1497 1519
FT CHAIN 1520 1738
FT CHAIN 1739 2227
FT VARIANT 77 77
FT VARIANT 764 764
FT VARIANT 821 821
FT VARIANT 1052 1052
FT VARIANT 1062 1062
FT VARIANT 1118 1118
FT VARIANT 1151 1151
FT VARIANT 1163 1163
FT VARIANT 1277 1277
FT VARIANT 1500 1500
FT VARIANT 1805 1805
FT VARIANT 1930 1930
SQ SEQUENCE 2227 AA; 251506 MW; 01E225E7AEB740A6 CRC64;
K -> R (IN ATTENUATED STRAIN).
E -> V (IN ATTENUATED STRAIN).
N -> S (IN ATTENUATED STRAIN).
A -> V (IN ATTENUATED STRAIN).
G -> A (IN ATTENUATED STRAIN).
K -> M (IN ATTENUATED STRAIN).
E -> K (IN ATTENUATED STRAIN).
F -> S (IN ATTENUATED STRAIN).
V -> I (IN ATTENUATED STRAIN).
H -> Y (IN ATTENUATED STRAIN).
D -> N (IN ATTENUATED STRAIN).
S -> T (IN ATTENUATED STRAIN).
RNA-DIRECTED POLYMERASE 3D.
K -> R (IN ATTENUATED STRAIN).
E -> V (IN ATTENUATED STRAIN).
N -> S (IN ATTENUATED STRAIN).
A -> V (IN ATTENUATED STRAIN).
G -> A (IN ATTENUATED STRAIN).
K -> M (IN ATTENUATED STRAIN).
E -> K (IN ATTENUATED STRAIN).
F -> S (IN ATTENUATED STRAIN).
V -> I (IN ATTENUATED STRAIN).
H -> Y (IN ATTENUATED STRAIN).
D -> N (IN ATTENUATED STRAIN).
S -> T (IN ATTENUATED STRAIN).
Query Match 100.0%; Score 102; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PYKELRLEVGVKQRLKYAQEE 20
Db 799 PYKELRLEVGVKQRLKYAQEE 818
|||||

RESULT 8
POLG_HPAVH STANDARD; PRT; 2227 AA.
AC P06441;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: COAT PROTEINS VP1 TO VP4; CORE PROTEINS
DE P2A TO P2C; PROBABLE PROTEINS P3A TO P3C; RNA-DIRECTED RNA POLYMERASE
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain LA).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12099;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85190549; PubMed=2986127;
RA Najarian K., Caput D., Gee W.W., Potter S.J., Renard A.,
RA Merryweather J., van Nest G., Dina D.;
RT "Primary structure and gene organization of human hepatitis A virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).
```

CC -|- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
 CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
 CC VP3, AND VP4.
 CC -|- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -|- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M20273; AAA45474.1; -
 CC PIR; JS0303; GNNYHB.
 CC DR MEROPS; C03.005; -
 CC DR InterPro; IPR000605; -
 CC DR InterPro; IPR001205; -
 CC DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 CC DR Pfam; PF00910; RNA_helicase; 1.
 CC KW Polyprotein; Coat protein; Core protein; Transferase;
 CC RNA-directed RNA polymerase; Hydrolase; Thiol protease;
 CC FT CHAIN 1 23
 CC FT CHAIN 24 245
 CC FT CHAIN 246 491
 CC FT CHAIN 492 836
 CC FT CHAIN 837 980
 CC FT CHAIN 981 1087
 CC FT CHAIN 1088 1422
 CC FT CHAIN 1423 1496
 CC FT CHAIN 1497 1519
 CC FT CHAIN 1520 1738
 CC FT CHAIN 1739 2227
 CC RNA-DIRECTED POLYMERASE P3D.
 CC SEQUENCE 2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;

Query Match 100.0%; Score 102; DB 1; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 4.8e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PYKELRLEVGKORLKYAQEE 20
 |||||
 DB 799 PYKELRLEVGKORLKYAQEE 818

RESULT 9
 POLG_HPVM STANDARD; PRT; 2227 AA.
 AC P13901; Q81083; Q81084; Q81085; Q81086; Q81087; Q81088; Q81089;
 AC Q81090; Q81091; Q81092; Q81093;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE GENOME POLYPROTEIN [CONTAINS: COAT PROTEINS VP1 TO VP4; CORE PROTEINS
 DE P2A TO P2C; PROBABLE PROTEINS P3A TO P3C; RNA-DIRECTED RNA POLYMERASE
 DE P3D (EC 2.7.7.48)].
 OS Hepatitis A virus (strain MBB).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12100;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88045071; PubMed=2823500;
 RA Paul A.V., Tada H., der Helm K., Wissel T., Kiehn R., Wimmer E.,
 RA Deinhardt F.;
 RT "The entire nucleotide sequence of the genome of human hepatitis A
 RT virus (isolate MBB).";
 RL Virus Res. 8:153-171(1987).
 CC -|- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
 CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
 CC VP3, AND VP4.
 CC -|- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -|- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT
 CC WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.
 CC -|- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M20273; AAA45474.1; -
 CC PIR; JS0303; GNNYHB.
 CC DR MEROPS; C03.005; -
 CC DR InterPro; IPR000605; -
 CC DR InterPro; IPR001205; -
 CC DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 CC DR Pfam; PF00910; RNA_helicase; 1.
 CC KW Polyprotein; Coat protein; Core protein; Transferase;
 CC RNA-directed RNA polymerase; Hydrolase; Thiol protease;
 CC FT CHAIN 1 23
 CC FT CHAIN 24 245
 CC FT CHAIN 246 491
 CC FT CHAIN 492 836
 CC FT CHAIN 837 980
 CC FT CHAIN 981 1087
 CC FT CHAIN 1088 1422
 CC FT CHAIN 1423 1496
 CC FT CHAIN 1497 1519
 CC FT CHAIN 1520 1738
 CC FT CHAIN 1739 2227
 CC RNA-DIRECTED POLYMERASE P3D.
 CC SEQUENCE 2227 AA; 251425 MW; BC983ED2A7C86349 CRC64;

Query Match 100.0%; Score 102; DB 1; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 4.8e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PYKELRLEVGKORLKYAQEE 20
 |||||
 DB 799 PYKELRLEVGKORLKYAQEE 818

RESULT 10
 POLG_HPVS STANDARD; PRT; 2230 AA.
 AC P14553;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE GENOME POLYPROTEIN [CONTAINS: COAT PROTEINS VP1 TO VP4; CORE PROTEINS
 DE P2A TO P2C; PROBABLE PROTEINS P3A TO P3C; RNA-DIRECTED RNA POLYMERASE
 DE P3D (EC 2.7.7.48)].
 OS Simian hepatitis A virus (strain AGM-27).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12102;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91311420; PubMed=1649901;
 RA Tsarev S.A., Emerson S.U., Balayan M.S., Ticehurst J.R.,
 RA Purcell R.H.;
 RT "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome
 RT structure and growth in cell culture with other HAV strains.";
 RL J. Gen. Virol. 72:1677-1683(1991).
 RN [2]
 RP SEQUENCE OF 1750-2164 FROM N.A.
 RX MEDLINE=89232168; PubMed=2541023;
 RA Balayan M.S., Kusov Y.Y., Andjapardze A.G., Tsarev S.A.,
 RA Sverdlov E.D., Chizhikov V.E., Blinov V.M., Vasilenko S.K.;
 RT "Variations in genome fragments coding for RNA polymerase in human
 RT and simian hepatitis A viruses.";

```

RL FEBS Lett. 247:425-428(1989).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS.
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D00924; BAA00766.1; -.
DR EMBL: X15461; CAA33490.1; -.
DR PIR: A30470; GNNYSA.
DR PIR: S04885; S04885.
DR MEROPS: C03.005; -.
DR InterPro: IPR000605; -.
DR InterPro: IPR001205; -.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 27 COAT PROTEIN VP4 (PIA).
FT CHAIN 28 249 COAT PROTEIN VP2 (PIB).
FT CHAIN 250 495 COAT PROTEIN VP3 (PIC).
FT CHAIN 496 795 COAT PROTEIN VP1 (PID).
FT CHAIN 796 984 CORE PROTEIN P2A.
FT CHAIN 985 1091 CORE PROTEIN P2B.
FT CHAIN 1092 1426 CORE PROTEIN P2C.
FT CHAIN 1427 1498 PROBABLE PROTEIN 3A.
FT CHAIN 1499 1521 PROBABLE PROTEIN 3B.
FT CHAIN 1522 1741 PROBABLE PROTEIN 3C.
FT CHAIN 1742 2230 RNA-DIRECTED POLYMERASE 3D.
SQ SEQUENCE 2230 AA; 251296 MW; 87B3230E324E1F19 CRC64;

Query Match 93.18; Score 95; DB 1; Length 2230;
Best Local Similarity 90.0%; Pred. No. 6.2e-07;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 PYKELRLEVGKQRLKYAEE 20
Db 803 PYKELRMEVGKQRLKYAEE 822

RESULT 11
POLG_HPAPT STANDARD; PRT; 839 AA.
AC P31788;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: COAT PROTEINS VP1 TO VP4; CORE PROTEIN
DE P2A] (FRAGMENT).
OS Simian hepatitis A virus (strain CY-145).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=31707;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91311421; PubMed=1649902;
RA Nainan O.V., Margolis H.S., Robertson B.H., Balayan M., Brinton M.A.;
RT "Sequence analysis of a new hepatitis A virus naturally infecting
RT cynomolgus macaques (Macaca fascicularis).";
RL J. Gen. Virol. 72:1685-1689(1991).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS.
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M59286; AAA45473.1; -.
DR PIR: J01180; GNNYS2.
KW Polyprotein; Coat protein; Core protein.
FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
FT CHAIN 492 ? COAT PROTEIN VP1 (PID).
FT CHAIN ? >839 CORE PROTEIN P2A.
FT NON_TER 839 839
SQ SEQUENCE 839 AA; 93825 MW; 2CACC4BD1E192DBC CRC64;

Query Match 92.2%; Score 94; DB 1; Length 839;
Best Local Similarity 90.0%; Pred. No. 3.1e-07;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 PYKELRLEVGKQRLKYAEE 20
Db 798 PYKELRLEVGKQRFKYAEE 817

RESULT 12
VATD_CHLPN STANDARD; PRT; 209 AA.
AC Q9Z991;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE V-TYPE ATP SYNTHASE SUBUNIT D (EC 3.6.1.34) (V-TYPE ATPASE SUBUNIT D).
GN ATPD OR CPN0090 OR CP0684.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Unayam L.A., Utterback T.,
RA Berry K., Bass S., Linher K., Weidman J., Khouri H., Craven B.,
RA Bowman C., Dodson R., Gwinn M., Nelson W., DeBoy R., Kolonay J.,
RA McClarty G., Salzberg S.L., Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shiba T., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
CC -!- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON
CC GRADIENT ACROSS THE MEMBRANE.

```

CC -!- SIMILARITY: BELONGS TO THE V-ATPASE D SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001594; AAD18243.1; -
DR EMBL: AE002227; AAF73696.1; -
DR EMBL: AP002545; BAA98300.1; -
DR TIGR: CP0684; -
DR InterPro: IPR002699; -
DR Pfam: PF01813; ATP-synt.D; 1.
KW Hydrolyase: ATP synthesis; Hydrogen ion transport.
SQ SEQUENCE 209 AA; 23788 MW; DA0A5C36313FE8BF CRC64;

Query Match 44.1%; Score 45; DB 1; Length 209;
Best Local Similarity 50.0%; Pred. No. 4;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Oy 3 KELLREVGKQRLKYAQEE 20
Db 133 KKVMAEVSKEKRLKILEEE 150

RESULT 13
ID ILVE_RICPR STANDARD; PRT; 290 AA.
AC 005970;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROBABLE BRANCHED-CHAIN AMINO ACID AMINOTRANSFERASE (EC 2.6.1.42)
DE (BCAT).
GN ILVE OR RP428.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsieae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MADRID E;
RX MEDLINE=97419517; PubMed=9274032;
RA Andersson J.O., Andersson S.G.E.;
RT "Genomic rearrangements during evolution of the obligate
RT intracellular parasite Rickettsia prowazekii as inferred from an
RT analysis of 52015 bp nucleotide sequence.";
RL Microbiology 143:2783-2795(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MADRID E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sichteritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria.";
RL Nature 396:133-140(1998).
CC -!- CATALYTIC ACTIVITY: L-LEUCINE + 2-OXOGLUTARATE = 4-METHYL-2-
CC OXOPENTANOATE + L-GLUTAMATE (ALSO ACTS ON L-ISOLEUCINE AND
CC L-VALINE).
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -!- PATHWAY: VALINE AND ISOLEUCINE BIOSYNTHESIS.
CC -!- SIMILARITY: BELONGS TO CLASS-IV OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y11777; CAA72450.1; -
DR EMBL: AJ235271; CAA14885.1; -
DR InterPro: IPR001544; -
DR Pfam: PF01063; aminotran_4; 1.
DR PROSITE: PS00770; AA-TRANSFER-CLASS.4; 1.
KW Transferase; Aminotransferase; Branched-chain amino acid biosynthesis;
KW Pyridoxal phosphate.
FT BINDING 155 155 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 290 AA; 33068 MW; 98374E325350763D CRC64;

Query Match 43.1%; Score 44; DB 1; Length 290;
Best Local Similarity 61.1%; Pred. No. 8.2;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Oy 3 KELLREVGKQRLKYAQEE 20
Db 225 KSLCLEVKEERLKLQAE 242

RESULT 14
FOJO_DROME
ID FOJO_DROME STANDARD; PRT; 583 AA.
AC P54360; O24176;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE FOUR-JOINTED PROTEIN.
GN FJ.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye Imaginal disk;
RX MEDLINE=96038089; PubMed=7555705;
RA Villano J.L., Katz F.N.;
RT "four-jointed is required for intermediate growth in the proximal-
RT distal axis in Drosophila";
RL Development 121:2767-2777(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CANTON-S;
RX MEDLINE=96187865; PubMed=8606003;
RA Brodsky M.H., Steller H.;
RT "Positional information along the dorsal-ventral axis of the
RT Drosophila eye: graded expression of the four-jointed gene.";
RL Dev. Biol. 173:428-446(1996).
CC -!- FUNCTION: REQUIRED FOR INTERMEDIATE GROWTH IN THE PROXIMAL-DISTAL
CC AXIS. MAY BE REQUIRED FOR CELL-CELL SIGNALING DURING DISC
CC DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN (POTENTIAL).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U28837; AAA69524.1; -
DR EMBL: U44904; AAB01809.1; -
DR FlyBase: FBgn0000658; fj.
KW Transmembrane; Signal-anchor.
FT DOMAIN 1 78 CYTOPLASMIC (POTENTIAL).
FT

Search completed: April 24, 2001, 16:33:55
Job time: 551 sec

```
FT TRANSMEM 79 99 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT DOMAIN 100 583 (POTENTIAL).
FT CARBOHYD 310 310 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 379 379 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 491 491 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 126 126 T -> S (IN REF. 2).
FT CONFLICT 193 193 M -> L (IN REF. 2).
FT CONFLICT 288 288 P -> R (IN REF. 2).
FT CONFLICT 330 330 R -> A (IN REF. 2).
SQ SEQUENCE 583 AA; 65504 MW; 82F1EA2A299DB284 CRC64;

Query Match 43.1%; Score 44; DB 1; Length 583;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 YKELRLEVKGQRLKYAQE 19
I III I: I: I:
DB 2 YDIKRLEAGQKQLQAQQ 19

RESULT 15
ID UPPE_AQUAE STANDARD; PRT; 208 AA.
AC 067914;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE URACIL PHOSPHORIBOSYLTRANSFERASE (EC 2.4.2.9) (UMP PYROPHOSPHORYLASE)
DE (UPRTASE).
DN UPP OR URAP OR AQ_2163.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Kellner M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus";
RL Nature 392:353-358(1998).
CC -!- CATALYTIC ACTIVITY: UMP + PYROPHOSPHATE = URACIL + 5-PHOSPHO-
CC ALPHA-D-RIBOSE 1-DIPHOSPHATE.
CC -!- PATHWAY: PYRIMIDINE SALVAGE PATHWAY.
CC -!- SIMILARITY: BELONGS TO THE UPRTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000776; AAC07880.1; -.
CC InterPro; IPR000836; -.
CC Pfam; PF00156; Pribosyltran; 1.
CC Transferase; Glycosyltransferase.
SQ SEQUENCE 208 AA; 23532 MW; 87B9CFEFA9CFF77 CRC64;

Query Match 42.2%; Score 43; DB 1; Length 208;
Best Local Similarity 44.4%; Pred. No. 8.3;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 KELRLEVKGQRLKYAQEE 20
I: I: I: I: I:
DB 51 KEVTRWIGNKRFNLYNEE 68
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2001, 16:33:07 ; Search time 113.97 Seconds
(without alignments)
20.568 Million cell updates/sec

Title: US-09-171-432A-40
Perfect score: 102
Sequence: 1 PYKELRLEVGKQRLKYAOEE 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_15:
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	102	100.0	56	14	Q67818 hepatitis a
2	102	100.0	56	14	Q67819 hepatitis a
3	102	100.0	56	14	Q67820 hepatitis a
4	102	100.0	56	14	Q67821 hepatitis a
5	102	100.0	56	14	Q67822 hepatitis a
6	102	100.0	56	14	Q67823 hepatitis a
7	102	100.0	56	14	Q39865 hepatitis a
8	102	100.0	56	14	Q39866 hepatitis a
9	102	100.0	56	14	Q39867 hepatitis a
10	102	100.0	56	14	Q39868 hepatitis a
11	102	100.0	56	14	Q39869 hepatitis a
12	102	100.0	56	14	Q39870 hepatitis a
13	102	100.0	56	14	Q39871 hepatitis a
14	102	100.0	56	14	Q39872 hepatitis a
15	102	100.0	56	14	Q39873 hepatitis a
16	102	100.0	56	14	Q39874 hepatitis a
17	102	100.0	56	14	Q96829 hepatitis a
18	102	100.0	56	14	Q96830 hepatitis a
19	102	100.0	56	14	Q91PD2 hepatitis a

20	102	100.0	56	14	Q91PD1 hepatitis a
21	102	100.0	56	14	Q91PC9 hepatitis a
22	102	100.0	56	14	Q91PC8 hepatitis a
23	102	100.0	56	14	Q91PC7 hepatitis a
24	102	100.0	56	14	Q91PC5 hepatitis a
25	102	100.0	56	14	Q91PC4 hepatitis a
26	102	100.0	56	14	Q91PC3 hepatitis a
27	102	100.0	56	14	Q91PC2 hepatitis a
28	102	100.0	56	14	Q91PC1 hepatitis a
29	102	100.0	56	14	Q91PC0 hepatitis a
30	102	100.0	56	14	Q91PB9 hepatitis a
31	102	100.0	56	14	Q91PB8 hepatitis a
32	102	100.0	56	14	Q91PB6 hepatitis a
33	102	100.0	56	14	Q91PB5 hepatitis a
34	102	100.0	56	14	Q91PB4 hepatitis a
35	102	100.0	56	14	Q91PB3 hepatitis a
36	102	100.0	56	14	Q91PB2 hepatitis a
37	102	100.0	56	14	Q91PB1 hepatitis a
38	102	100.0	56	14	Q91PB0 hepatitis a
39	102	100.0	56	14	Q91PA9 hepatitis a
40	102	100.0	56	14	Q91PA8 hepatitis a
41	102	100.0	56	14	Q91PA7 hepatitis a
42	102	100.0	56	14	Q91PA6 hepatitis a
43	102	100.0	56	14	Q91PA5 hepatitis a
44	102	100.0	56	14	Q91PA4 hepatitis a
45	102	100.0	56	14	Q91PB7 hepatitis a

ALIGNMENTS

RESULT 1
Q67818
ID Q67818 PRELIMINARY; PRT; 56 AA.
AC Q67818;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TremBLrel. 13, Last annotation update)
DE POLYPROTEIN (FRAGMENTS).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
RX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISR-05, AFULA DISTRICT, ISRAEL, 1993;
RA Kartnnyi Y.V., Shulman L.M., Manor J., Leitner L., Shehab S.,
RA Mendelson E.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; 277243; CAB01036.1; -;
DR INTERPRO; IPR000886; -;
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW Polyprotein; Nonstructural protein.
FT NON_TER 1
FT CHAIN 1 >28 CAPSID PROTEIN VP1.
FT NON_CONS 28 29
FT CHAIN 29 >56 NONSTRUCTURAL PROTEIN NS2A.
FT NON_TER 56 56
SQ SEQUENCE 56 AA: 6614 MW: 843851846AEF4A4 CRC64;

Query Match 100.0%; Score 102; DB 14; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAOEE 20
|||||
Db 36 PYKELRLEVGKQRLKYAOEE 55

RESULT 2
Q67819
ID Q67819 PRELIMINARY; PRT; 56 AA.

```
AC Q67819;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE POLYPROTEIN (FRAGMENTS).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISR-10, (AFULA DISTRICT, ISRAEL, 1993);
RA Karetnyi Y.V., Shulman L.M., Manor J., Leitner L., Shehab S.,
RA Mendelson E.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z77244; CAB01037.1; -.
DR INTERPRO; IPR000886; -.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW Polyprotein; Nonstructural protein.
FT NON_TER 1 1
FT CHAIN 1 >28 CAPSID PROTEIN VP1.
FT NON_CONS 28 29
FT CHAIN 29 >56 NONSTRUCTURAL PROTEIN NS2A.
FT NON_TER 56 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 102; DB 14; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
DB 36 PYKELRLEVGKQRLKYAQEE 55

RESULT 3
ID Q67820 PRELIMINARY; PRT; 56 AA.
AC Q67820;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE POLYPROTEIN (FRAGMENTS).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISR-12, (AFULA DISTRICT, ISRAEL, 1993);
RA Karetnyi Y.V., Shulman L.M., Manor J., Leitner L., Shehab S.,
RA Mendelson E.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z77245; CAB01038.1; -.
DR INTERPRO; IPR000886; -.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW Polyprotein; Nonstructural protein.
FT NON_TER 1 1
FT CHAIN 1 >28 CAPSID PROTEIN VP1.
FT NON_CONS 28 29
FT CHAIN 29 >56 NONSTRUCTURAL PROTEIN NS2A.
FT NON_TER 56 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 102; DB 14; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
DB 36 PYKELRLEVGKQRLKYAQEE 55

RESULT 4
ID Q67821 PRELIMINARY; PRT; 56 AA.
AC Q67821;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE POLYPROTEIN (FRAGMENTS).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISR-15, (AFULA DISTRICT, ISRAEL, 1993);
RA Karetnyi Y.V., Shulman L.M., Manor J., Leitner L., Shehab S.,
RA Mendelson E.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z77246; CAB01039.1; -.
DR INTERPRO; IPR000886; -.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW Polyprotein; Nonstructural protein.
FT NON_TER 1 1
FT CHAIN 1 >28 CAPSID PROTEIN VP1.
FT NON_CONS 28 29
FT CHAIN 29 >56 NONSTRUCTURAL PROTEIN NS2A.
FT NON_TER 56 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 102; DB 14; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
DB 36 PYKELRLEVGKQRLKYAQEE 55

RESULT 5
ID Q67822 PRELIMINARY; PRT; 56 AA.
AC Q67822;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE POLYPROTEIN (FRAGMENTS).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISR-70, (AKKO DISTRICT, ISRAEL, 1993);
RA Karetnyi Y.V., Shulman L.M., Manor J., Leitner L., Shehab S.,
RA Mendelson E.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z77247; CAB01040.1; -.
DR INTERPRO; IPR000886; -.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW Polyprotein; Nonstructural protein.
FT NON_TER 1 1
FT CHAIN 1 >28 CAPSID PROTEIN VP1.
FT NON_CONS 28 29
FT CHAIN 29 >56 NONSTRUCTURAL PROTEIN NS2A.
FT NON_TER 56 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 102; DB 14; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
DB 36 PYKELRLEVGKQRLKYAQEE 55
```

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAEE 20
 Db 36 PYKELRLEVGKQRLKYAEE 55

RESULT 6

ID Q67823 PRELIMINARY; PRT; 56 AA.
 AC Q67823;
 DT 01-NOV-1996 (T-EMBLrel. 01, Created)
 DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
 DT 01-MAY-2000 (T-EMBLrel. 13, Last annotation update)
 DE POLYPROTEIN (FRAGMENTS).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ISR-79, (AKKO DISTRICT, ISRAEL, 1993);
 RA Karetnyi Y.V., Shulman L.M., Manor J., Leitner L., Shehab S.,
 RA Hendelson E.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; 277248; CAB01041.1; -;
 DR INTERPRO; IPR000886; -;
 DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
 KW Polyprotein; Nonstructural protein.
 FT NON_TER 1
 FT CHAIN 1 >28 CAPSID PROTEIN VP1.
 FT NON_CONS 28 29
 FT CHAIN 29 >56 NONSTRUCTURAL PROTEIN NS2A.
 FT NON_TER 56 56
 FT SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 102; DB 14; Length 56;
 Best Local Similarity 100.0%; Pred. No. 2.9e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAEE 20
 Db 36 PYKELRLEVGKQRLKYAEE 55

RESULT 7

ID O39865 PRELIMINARY; PRT; 56 AA.
 AC O39865;
 DT 01-JAN-1998 (T-EMBLrel. 05, Created)
 DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
 DT 01-MAY-2000 (T-EMBLrel. 13, Last annotation update)
 DE POLYPROTEIN (FRAGMENT).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=412991;
 RX MEDLINE=97247817; PubMed-9093940;
 RA Taylor M.B.;
 RT "Molecular epidemiology of South African strains of hepatitis A virus;
 1982-1996.";
 RL J. Med. Virol. 51:273-279(1997).
 DR EMBL; U68689; AAB53585.1; -;
 DR INTERPRO; IPR000886; -;
 DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
 KW Polyprotein.
 FT NON_TER 1
 FT NON_TER 56 56
 FT SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 102; DB 14; Length 56;
 Best Local Similarity 100.0%; Pred. No. 2.9e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAEE 20
 Db 36 PYKELRLEVGKQRLKYAEE 55

RESULT 8

ID O39866 PRELIMINARY; PRT; 56 AA.
 AC O39866;
 DT 01-JAN-1998 (T-EMBLrel. 05, Created)
 DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
 DT 01-MAY-2000 (T-EMBLrel. 13, Last annotation update)
 DE POLYPROTEIN (FRAGMENT).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=96001190;
 RX MEDLINE=97247817; PubMed-9093940;
 RA Taylor M.B.;
 RT "Molecular epidemiology of South African strains of hepatitis A virus;
 1982-1996.";
 RL J. Med. Virol. 51:273-279(1997).
 DR EMBL; U68690; AAB53586.1; -;
 DR INTERPRO; IPR000886; -;
 DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
 KW Polyprotein.
 FT NON_TER 1
 FT NON_TER 56 56
 FT SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 102; DB 14; Length 56;
 Best Local Similarity 100.0%; Pred. No. 2.9e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAEE 20
 Db 36 PYKELRLEVGKQRLKYAEE 55

RESULT 9

ID O39867 PRELIMINARY; PRT; 56 AA.
 AC O39867;
 DT 01-JAN-1998 (T-EMBLrel. 05, Created)
 DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
 DT 01-MAY-2000 (T-EMBLrel. 13, Last annotation update)
 DE POLYPROTEIN (FRAGMENT).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JVR;
 RX MEDLINE=97247817; PubMed-9093940;
 RA Taylor M.B.;
 RT "Molecular epidemiology of South African strains of hepatitis A virus;
 1982-1996.";
 RL J. Med. Virol. 51:273-279(1997).
 DR EMBL; U68692; AAB53588.1; -;
 DR INTERPRO; IPR000886; -;
 DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
 KW Polyprotein.

```
FT  NON_TER      1      1
FT  NON_TER     56     56
SQ  SEQUENCE    56 AA;  8438C51846AEF4A4 CRC64;

Query Match      100.0%; Score 102; DB 14; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 PYKELRLEVKGKORLYAQEE 20
    |||
Db  36 PYKELRLEVKGKORLYAQEE 55
    |||

RESULT 10
O39868
ID  O39868      PRELIMINARY;      PRT;      56 AA.
AC  O39868;
DT  01-JAN-1998 (TrEMBLrel. 05, Created)
DT  01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT  01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE  POLYPROTEIN (FRAGMENT).
OS  Hepatitis A virus.
OC  Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC  Hepatovirus.
OX  NCBI_TaxID=12092;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=503712;
RX  MEDLINE=97247817; PubMed=9093940;
RA  Taylor M.B.;
RT  "Molecular epidemiology of South African strains of hepatitis A virus:
    J. Med. Virol. 51:273-279(1997).
DR  EMBL; U68693; AAB53589.1; -.
DR  INTERPRO; IPR000886; -.
DR  PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW  Polyprotein.
FT  NON_TER      1      1
FT  NON_TER     56     56
SQ  SEQUENCE    56 AA;  8438C51846AEF4A4 CRC64;

Query Match      100.0%; Score 102; DB 14; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 PYKELRLEVKGKORLYAQEE 20
    |||
Db  36 PYKELRLEVKGKORLYAQEE 55
    |||

RESULT 11
O39869
ID  O39869      PRELIMINARY;      PRT;      56 AA.
AC  O39869;
DT  01-JAN-1998 (TrEMBLrel. 05, Created)
DT  01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT  01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE  POLYPROTEIN (FRAGMENT).
OS  Hepatitis A virus.
OC  Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC  Hepatovirus.
OX  NCBI_TaxID=12092;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=96002382;
RX  MEDLINE=97247817; PubMed=9093940;
RA  Taylor M.B.;
RT  "Molecular epidemiology of South African strains of hepatitis A virus:
    J. Med. Virol. 51:273-279(1997).
DR  EMBL; U68694; AAB53590.1; -.

DR  INTERPRO; IPR000886; -.
DR  PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW  Polyprotein.
FT  NON_TER      1      1
FT  NON_TER     56     56
SQ  SEQUENCE    56 AA;  8438C51846AEF4A4 CRC64;

Query Match      100.0%; Score 102; DB 14; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 PYKELRLEVKGKORLYAQEE 20
    |||
Db  36 PYKELRLEVKGKORLYAQEE 55
    |||

RESULT 12
O39870
ID  O39870      PRELIMINARY;      PRT;      56 AA.
AC  O39870;
DT  01-JAN-1998 (TrEMBLrel. 05, Created)
DT  01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT  01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE  POLYPROTEIN (FRAGMENT).
OS  Hepatitis A virus.
OC  Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC  Hepatovirus.
OX  NCBI_TaxID=12092;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=406808;
RX  MEDLINE=97247817; PubMed=9093940;
RA  Taylor M.B.;
RT  "Molecular epidemiology of South African strains of hepatitis A virus:
    J. Med. Virol. 51:273-279(1997).
DR  EMBL; U68695; AAB53591.1; -.
DR  INTERPRO; IPR000886; -.
DR  PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW  Polyprotein.
FT  NON_TER      1      1
FT  NON_TER     56     56
SQ  SEQUENCE    56 AA;  6642 MW;  465CF51846AEF4BC CRC64;

Query Match      100.0%; Score 102; DB 14; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 PYKELRLEVKGKORLYAQEE 20
    |||
Db  36 PYKELRLEVKGKORLYAQEE 55
    |||

RESULT 13
O39871
ID  O39871      PRELIMINARY;      PRT;      56 AA.
AC  O39871;
DT  01-JAN-1998 (TrEMBLrel. 05, Created)
DT  01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT  01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE  POLYPROTEIN (FRAGMENT).
OS  Hepatitis A virus.
OC  Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC  Hepatovirus.
OX  NCBI_TaxID=12092;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=923359;
RX  MEDLINE=97247817; PubMed=9093940;
RA  Taylor M.B.;
RT  "Molecular epidemiology of South African strains of hepatitis A virus:
    J. Med. Virol. 51:273-279(1997).
DR  EMBL; U68694; AAB53590.1; -.
DR  INTERPRO; IPR000886; -.
DR  PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW  Polyprotein.
FT  NON_TER      1      1
FT  NON_TER     56     56
SQ  SEQUENCE    56 AA;  6614 MW;  8438C51846AEF4A4 CRC64;

Query Match      100.0%; Score 102; DB 14; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 PYKELRLEVKGKORLYAQEE 20
    |||
Db  36 PYKELRLEVKGKORLYAQEE 55
    |||
```

```

RT 1982-1996.*;
RL J. Med. Virol. 51:273-279(1997).
DR EMBL; U68696; AAB53592.1; -.
DR INTERPRO; IPR000886; -.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;

Query Match 100.0%; Score 102; DB 14; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
DB 36 PYKELRLEVGKQRLKYAQEE 55

RESULT 14
O39872
ID O39872 PRELIMINARY; PRT; 56 AA.
AC O39872;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
DE POLYPROTEIN (FRAGMENT).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VDM;
RX MEDLINE=97247817; PubMed=9093940;
RA Taylor M.B.;
RT 1982-1996.*;
RL J. Med. Virol. 51:273-279(1997).
DR EMBL; U68697; AAB53593.1; -.
DR INTERPRO; IPR000886; -.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;

Query Match 100.0%; Score 102; DB 14; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
DB 36 PYKELRLEVGKQRLKYAQEE 55

RESULT 15
O39873
ID O39873 PRELIMINARY; PRT; 56 AA.
AC O39873;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
DE POLYPROTEIN (FRAGMENT).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2333;
```

```

RX MEDLINE=97247817; PubMed=9093940;
RA Taylor M.B.;
RT "Molecular epidemiology of South African strains of hepatitis A virus:
  1982-1996.*";
RL J. Med. Virol. 51:273-279(1997).
DR EMBL; U68698; AAB53594.1; -.
DR INTERPRO; IPR000886; -.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 102; DB 14; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
DB 36 PYKELRLEVGKQRLKYAQEE 55
```

Search completed: April 24, 2001, 16:33:07
Job time: 568 sec

